Online Advertisement Allocation Under Customer Choices and Algorithmic Fairness

Xiaolong Li
National University of Singapore, oralxi@nus.edu.sg

Ying Rong
Shanghai Jiao Tong University, yrong@sjtu.edu.cn

Renyu Zhang
New York University Shanghai, renyu.zhang@nyu.edu

Huan Zheng
Shanghai Jiao Tong University, zhenghuan@sjtu.edu.cn

December 12, 2021

Advertising is a major revenue source for e-commerce platforms and an important online marketing tool for e-commerce sellers. In this paper, we explore dynamic ad allocation with limited slots upon each customer arrival for e-commerce platforms when customers follow a choice model to click the ads. Motivated by the recent advocacy for the algorithmic fairness of online ad delivery, we adjust the value from advertising by a general fairness metric evaluated with the click-throughs of different ads and customer types. The original online ad-allocation problem is intractable, so we propose a novel, two-stage stochastic program reformulation that first decides the click-through targets then devises an ad-allocation policy to satisfy these targets in the second stage. We show the asymptotic equivalence between the original problem, the relaxed click-through target optimization, and the fluid-approximation (FA) convex program. We also design a debt-weighted offer-set (DWO) algorithm and demonstrate that, as long as the problem size scales to infinity, this algorithm is (asymptotically) optimal under the optimal first-stage click-through target. Compared to the FA heuristic and its re-solving variants, our approach has better scalability and can deplete the ad budgets more smoothly throughout the horizon, which is highly desirable for the online advertising business in practice. Finally, our proposed model and algorithm help substantially improve the fairness of ad allocation for an online e-commerce platform without compromising its efficiency much.

Key words: Online Advertising Platform, Choice Models, Algorithmic Fairness, Online Convex Optimization

---

1. Introduction

The past 10 years have witnessed a rapid growth of internet technology and smartphone penetration, which have driven online advertising to become an unprecedentedly enormous trillion-dollar industry[1] that has an enormous impact on the entire economy.

[1] See the report from the Interactive Advertising Bureau at https://www.iab.com/insights/internet-advertising-revenue-fy2019-q12020/ for more details.

One important online advertising format is e-commerce advertising, a form of online marketing or advertising that drives ”top-of-tunnel” traffic to convert into product sales. For instance, Amazon Advertising provides “sponsored products,”[2] where advertisers pay Amazon to promote their products by listing the ads both within the shopping results and on the product pages (see, Figure 1). The sponsored-product ads use the cost-per-click (CPC) mechanism, under which advertisers pay a fee to the platform when customers click their ads. Advertisers choose the campaign budgets and how much to bid per click for their ads. Amazon also allows advertisers to set the keywords and products so that the ad can be more efficiently matched with customer queries. Alternatively, advertisers can select automatic targeting to allow Amazon to match their ads to relevant search terms and products. This advertising service is an important source of revenue for Amazon, primarily included in service offerings that contributed US$14.1 billion to (5.02% of) its *annual net sales* in 2019.[3] As another example, Facebook launched its Dynamic Ads service to promote advertisers’ products to people who expressed an interest in relevant keywords or similar products.[4] The Dynamic Ads will automatically choose products from the catalog provided by the advertisers and display them to customers.

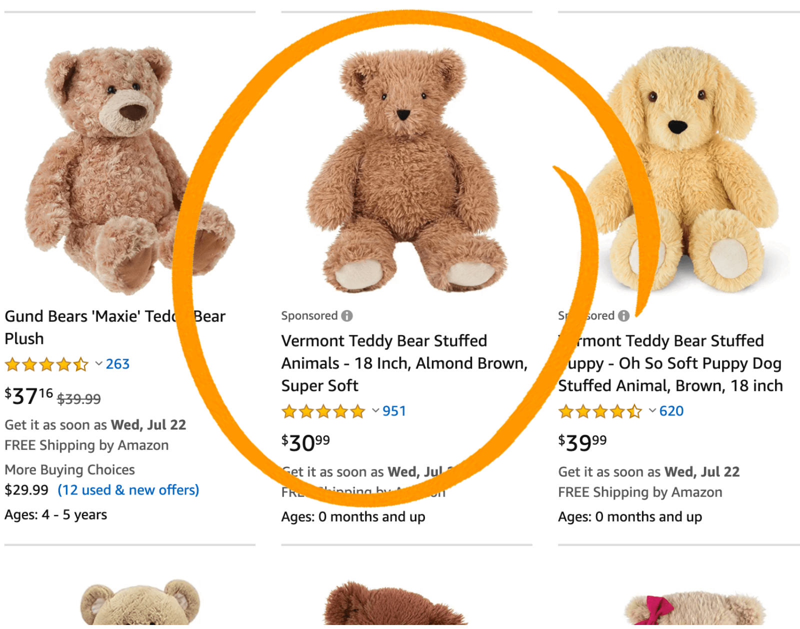


Figure 1 An Example of Sponsored Products on Amazon

Such large-scale e-commerce advertising platforms generally run thousands of advertising campaigns for different advertisers simultaneously. Each campaign is usually associated with (1) a budget

[2] https://advertising.amazon.com/solutions/products/sponsored-products.

[3] See Amazon’s 2019 financial report: https://www.sec.gov/ix?doc=/Archives/edgar/data/1018724/000101872420000004/amzn-20191231x10k.htm.

[4] See https://www.facebook.com/business/help/397103717129942?id=1913105122334058 for more details.

that the advertiser wishes to spend as much as possible of during the campaign horizon and (2) a bid per CPC that dictates how much ad budget should be deducted upon each user click. The advertising platform dynamically allocates its ad spaces (i.e., customer impressions) to the ads whose campaigns are active. As discussed above, an advertiser may require the platform to target his/her advertising campaign and ads to specific customer segments (specified by such features as location, age, and browsing, searching, and purchasing histories). It is also not uncommon for advertisers and, thus, the platform to set click-through requirements for the ads (i.e., the minimal number of click-throughs during the entire ad campaign). For instance, Microsoft provides a Partner Incentive Cooperative Marketing Fund (Co-op) to subsidize its partners in whose website the number of click-throughs for Microsoft's ad is above 250 during the promotion events (Microsoft 2020). In addition, from a long-term perspective, the number of click-throughs for an ad has a substantial impact on the long-term retention of the advertiser, which prompts the platform to devise the ad-allocation policy to secure a certain number of click-throughs for each ad (see, e.g., Ye et al. 2020).

To efficiently allocate its ad spaces, an online advertising platform faces the central operations problem of dynamically selecting a set of ads, which we refer to as an offer-set, displayed to each arriving customer in order to generate the highest total value throughout the planning horizon. In addition to optimizing the total advertising value subject to the budget and click-through requirement constraints, the online platform also needs to address the fairness/discrimination concerns of its advertising/machine-learning algorithms. The algorithmic fairness/discrimination concerns have been receiving increased attention recently in a broad range of literature spanning computer science, operations management, and business strategy (see, e.g., Barocas et al. 2017, Ali et al. 2019, Lambrecht and Tucker 2019, Lejeune and Turner 2019, Kallus et al. 2021). It has been well-documented in the literature that a common source of algorithmic discrimination/bias in online advertising is that advertisers can target (or exclude) particular groups of users for their ads (e.g., Speicher et al. 2018).

To counter the algorithmic discrimination/bias, large-scale online advertising platforms such as Facebook and Google have tightened their controls recently to prevent clients from excluding some user groups from seeing job, housing, and other ads, in order to reduce lawsuits and regulatory probes on discrimination through ad targeting.[5] However, simply banning advertisers from targeting the customer groups in a discriminatory fashion is insufficient to prevent discrimination/bias and restore fairness for online advertising. This is because the platform's underlying ad-allocation algorithm for optimizing certain business objectives, such as advertising revenue or advertisers' return on investment, may automatically skew ad delivery to certain user groups (see, e.g., Speicher et al. 2018, Lambrecht and Tucker 2019, Imana et al. 2021).

[5] See `https://www.reuters.com/technology/study-flags-gender-bias-facebooks-ads-tools-2021-04-09/` for more details.

The main goal of this paper is to explore the e-commerce ad allocation of an advertising platform under customer choices and concerns for algorithmic fairness. Motivated by online e-commerce advertising practices, we seek to address the following key research question:

> *Taking into account algorithmic fairness, how should a platform dynamically personalize the ad offer-sets of each customer impression to maximize the fairness-adjusted value (FAV)from advertising throughout a planning horizon in the presence of advertising budgets and click-through requirements for different ads?*

Answering this question presents a multifaceted challenge. First, under the customer-choice behavior when an offer-set is displayed, the platform has to carefully balance the notorious trade-off in assortment optimization between expanding the offer-set to enlarge the market share and keeping it small enough to reduce the cannibalization between different ads. Second, algorithmic fairness is typically measured based on the overall ad allocation of the entire campaign. Thus, unlike the advertising value collected upon each customer click, it is challenging to decompose the algorithmic fairness by time and optimize it in an adaptive fashion. Finally, the click-through requirements, budget constraints, and advertisers' targeting rules altogether raise the difficulty of even searching for a feasible (but not necessarily optimal) ad allocation.

In this paper, we present a general stochastic program model to study this complex dynamic ad-allocation problem and propose a scalable algorithm that yields the (asymptotically) optimal ad allocation. One salient feature of our model is the explicit incorporation of a general fairness metric that measures the discrimination/bias of the ad-allocation policy throughout the planning horizon. This feature will prove useful to gear the ad-delivery algorithm toward a desirable outcome that enjoys fairness and impartiality, thus providing a new feasible approach to address the aforementioned algorithmic discrimination and bias in ad delivery optimization.

Though the original ad-allocation problem is intractable, we *reformulate/relax* it into a two-stage stochastic program: in the first stage, the platform optimizes the click-through targets for each ad/customer-type pair during the entire campaign; in the second stage, it devises the ad offer-set policy to satisfy these targets. We show the (asymptotic) equivalence between the original stochastic program, the relaxed click-through target optimization, and the standard FA convex program. Based on this equivalence result, we design a simple, scalable DWO policy to satisfy the optimal click-through targets and achieve the asymptotically optimal FAV from advertising. We also conduct extensive numerical experiments to uncover several interesting insights on—relative to the commonly used benchmarks in the literature—the benefits of our proposed algorithm to improve both the efficiency and the fairness of an online advertising system.

1.1. Main Contributions

This paper's contributions can be summarized as follows:

Two-Stage Stochastic Program Reformulation. Given the intractability of the original online ad allocation for an e-commerce platform, we develop a novel (nonequivalent) two-stage stochastic program reformulation to study this problem. In the first-stage, the platform decides the click-through goal for each ad/customer-type pair to maximize the expected FAV, which is essentially reduced to a tractable, deterministic convex program. This reformulation provides a new upper bound for the original stochastic program for ad-allocation optimization. In the second-stage, an ad offer-set is displayed to each user upon her arrival in order to satisfy the optimal click-through targets set in the first stage. Our approach also enables us to characterize the necessary and sufficient condition under which the click-through targets are feasible. The characterization of feasible click-through targets is tractable for most commonly used choice models: multinomial logit, generalized attraction, and independent choices.

DWO **Algorithm.** As our main contribution here, we propose a simple and effective algorithm, referred to as the *debt-weighted offer-set policy*, and we demonstrate its (asymptotic) optimality. This policy dynamically assigns a "debt" to each click-through target that measures the difference between the realized total click-throughs and the endogenous target set in the first-stage convex program. Then, a standard offer-set/assortment optimization problem is solved to maximize a debt-weighted value function upon the arrival of each customer. By reweighting the ad values with the debts associated with a feasible click-through target vector, our proposed DWO algorithm generates an offer-set policy that *exactly* meets the (feasible) target set in the first-stage of the two-stage stochastic program.

Whereas most existing debt-weighted algorithms (e.g., Zhong et al. 2017, Jiang et al. 2019) focus on satisfying the feasibility/approchability of certain service-level constraints, we provide the *optimality* guarantee of the DWO policy. We also offer an interesting technological contribution to the literature with our proof of such optimality. Leveraging the *exact-approchability* result above, we establish new intrinsic connections and asymptotic equivalence of the original ad-allocation problem, the first-stage convex program, and an auxiliary FA convex program, implying that the DWO policy can achieve the theoretical upper bound of fairness-adjusted value characterized by the first-stage convex program. If the first-stage click-through target vector is only feasible, but not optimal, the associated DWO policy will not incur any additional optimality loss on top of that from the suboptimal target.

The DWO policy proposed in this paper is computationally scalable if customer choices follow commonly adopted models such as multinomial logit, generalized attraction, and independent choices. Through numerical experiments, we show that our algorithm performs better, more robustly, and more computationally efficiently than the commonly adopted FA-based benchmarks in the literature (e.g., Liu and Van Ryzin 2008, Jasin and Kumar 2012, Bumpensanti and Wang 2020) for most problem instances. Our numerical experiments also demonstrate that the proposed DWO algorithm provides much smoother depletion of budgets over the entire planning horizon than the benchmarks. This highlights the practical applicability of our approach, because smooth budget depletion is a very desirable property for the real-world online advertising business. Importantly, our model and the DWO algorithm help substantially improve the algorithmic fairness of ad allocation for an online e-commerce platform without compromising its efficiency much.

In summary, the key takeaway from this paper is that the proposed two-stage stochastic program reformulation together with the associated debt-weighted off-set algorithm can efficiently address the ad-allocation problem for e-commerce platforms to improve the FAV from advertising. Our approach is simple, efficient, and scalable, with a provable optimality guarantee and strong numerical performances. The rest of this paper is organized as follows. We review related literature in Section 2. We introduce the model in Section 3, and we propose the two-stage stochastic program formulation in Section 4. We study the optimal ad-allocation policy in Section 5, and we present the numerical studies in Section 6. Section 7 concludes. All proofs are relegated to the Online Appendices. Throughout this paper, we use a tilde to represent random variables/vectors, **boldface** to represent vectors and matrices, and $\boldsymbol{a} \geq \boldsymbol{b}$ to represent $a_j \geq b_j$ for each j.

2. Literature Review

This paper proposes a general modeling framework and efficient algorithms to study optimal online ad allocation for an e-commerce platform under algorithmic-fairness concerns. Our paper is primarily related to four streams of research in the literature: (a) ad allocation for online advertising platforms, (b) algorithmic discrimination/bias in online advertising, (c) resource allocation with individualized service-level constraints, and (d) (dynamic) personalized assortment optimization. Papers in the literature generally focus on some perspective of the four topics above, whereas our work contributes to all four streams of literature jointly.

Ad allocation is a central challenge for online advertising platforms. Scheduling advertisement display on websites has been widely studied in the literature. For example, Nakamura and Abe (2005) propose an ad-targeting approach based on linear programming that achieves high click-through rates by optimizing ad-display probabilities. Furthermore, a queue is adopted to hold the selected ads to avoid displaying the same ad more than once on one page and to reduce the display probabilities with small margins. Yang et al. (2012) combine an ad-inventory allocation problem with multi-objective of revenue and fairness. For maximizing reach of customers and minimizing variance of the outcome simultaneously in targeted advertising, Turner (2012) formulates an-planning problem with a quadratic objective to spread ads across all targeted customer types. Balseiro et al. (2014) formalize an ad-exchange problem as a multi-objective stochastic control problem considering both the revenue from exchange and click-through rates, and derive an efficient policy for online ad allocation with uncertainty. Also for dealing with uncertainty, Shen et al. (2021a) propose an integrated planning model with a distributionally robust chance-constrained program in online ad allocation. Hojjat et al. (2017) consider a new contract to allow advertisers to specify the number of unique individuals who should see their ad and the minimum number of times each individual should be exposed. They also introduce a new mechanism for ad serving that "pre-generates" an explicit sequence of ads for each user to see over time. Shen et al. (2021b) deal with customers' ad-clicking behavior by an arbitrary-point-inflated Poisson regression model, and they solve a mixed-integer nonlinear programming model for optimal ad allocation. We refer interested readers to Choi et al. (2020) for a comprehensive review of this literature. The key modeling difference of our paper from this literature is that, using choice models, we clearly model the click-through behaviors of a customer in the presence of multiple ads displayed to her simultaneously.

As mentioned earlier, algorithmic discrimination/bias in online advertising has received increased scrutiny in recent literature. Several works find evidence of algorithmic discrimination/bias based on race or gender in practical online advertising platforms. Speicher et al. (2018) show that simply disallowing the use of some sensitive features, such as race and gender, in targeted advertising is not sufficient to remove algorithmic bias. By exploring the data from a field test, Lambrecht and Tucker (2019) find that an algorithm delivered fewer job-opportunity ads to women than to men because it was more expensive to deliver the ads to young women. Though the algorithm optimizes ad delivery to achieve cost-effectiveness in a gender-neutral way, its intrinsic mechanism autonomously leads to bias and discrimination. Ali et al. (2019) find that, despite the neutral targeting parameters, both the budget and the content of the ad significantly contribute to Facebook's ad-delivery bias along gender and racial lines for ads on employment and housing opportunities. Imana et al. (2021) develop a new methodology for identifying discrimination due to protected categories such as gender and race in the delivery of job advertisements, distinguish this discrimination from legal bias due to differences in qualification among users, and confirm this discrimination by gender in ad delivery on Facebook. To mitigate the algorithmic discrimination/bias in online advertising, Celis et al. (2019) present a constrained ad-auction framework that maximizes the platform revenue subject to the condition that the number of users seeing an ad is distributed appropriately across sensitive user types such as gender and race. They solve a nonconvex optimization problem to obtain the optimal auction mechanism for a large class of fairness constraints. Lejeune and Turner (2019) derive a Gini-index-based metric to measure how well dispersed the impressions are allocated across audience segments, and they formulate an optimization problem to maximize the spread of impressions across targeted audience segments while minimizing demand shortfalls. In this paper, we offer a new slant on mitigating algorithmic discrimination/bias in online advertising. Specifically, we study a dynamic ad-allocation problem to maximize the FAV from advertising, whereas the models studied in the literature are mostly static. We propose a framework for capturing the efficiency-fairness trade-off in dynamic ad delivery, and we design a novel debt-based policy that balances this trade-off.

The resource-allocation problem of meeting service-target constraints in the face of uncertain demand has been extensively studied in the inventory literature (see, e.g., Eppen 1979, Swaminathan and Srinivasan 1999, Alptekinoğlu et al. 2013). Leveraging Blackwell's approachability theorem, Zhong et al. (2017) characterize the optimal safety-stock level with individual type-II service-level constraints. Lyu et al. (2019) and Lyu et al. (2021), respectively, extend both the approach and the results to the context of type-I service-level constraints and process flexibility. Utilizing a semi-infinite linear program formulation, Jiang et al. (2019) generalize and unify models in this literature and propose a simple randomized rationing policy to meet general service-level constraints, including type-I and type-II constraints, and beyond. Ma and Xu (2020) consider an online-matching problem with concerns of agent-group fairness; they define two different service-level objectives, instead of maximizing the number of matches, as the metrics of long-run and short-run fairness; and they show competitive ratios of their algorithms. Our contribution to this strand of the literature is that we generalize the concept of service-level constraints to incorporate customer choice uncertainty and indirect resource allocation through assortment planning. We also propose a family of debt-weighted offer-set algorithms and demonstrate their optimality for meeting the endogenous service targets and for generating the total payoff for the platform.

Over the past ten years, online e-commerce platforms have typically provided numerous products for customers to choose from (Feldman et al. 2021). Manufacturing firms have also expanded their product lines due to business trends (e.g., fast fashion, Caro et al. 2014) or technology revolution (e.g., 3D printing, Dong et al. 2017). The ever-expanding product pool makes personalized assortments more attractive. Therefore, personalized-assortment optimization has also been receiving increased attention in the literature. Leveraging the competitive-ratio framework, Golrezaei et al. (2014) propose inventory-balancing algorithms that guarantee the worst-case revenue performance without any forecast of the customer type distribution. Bernstein et al. (2019) combine dynamic assortment planning, demand learning, and customer-type clustering in a Bayesian framework and propose a prescriptive assortment-personalization approach for online retailing. Using re-solving heuristics, Jasin and Kumar (2012) study dynamic personalized assortment planning in a network revenue-management framework and show that the proposed heuristics achieve a constant optimality loss. Bumpensanti and Wang (2020) design a new, infrequent re-solving heuristic which re-solves the deterministic linear program less frequently at the beginning of the planning horizon and more frequently at the end. This infrequent re-solving heuristic achieves a revenue loss uniformly bounded by a constant independent of the time-horizon length and resource capacities. Kallus and Udell (2020) consider a dynamic assortment-personalization problem with a large number of items and customer types as a discrete contextual-bandit problem and propose a structural approach with efficient optimization algorithms. Chen et al. (2016) formalize a new checkout-recommender system at Walmart's online grocery as an online assortment-optimization problem with limited inventory, and they propose an inventory-protection algorithm with a bounded competitive ratio. A general, personalized resource-allocation model with customer choices is studied by Gallego et al. (2016). Adopting the column-generation approach to solve the choice-based linear program, the authors introduce algorithms with theoretical performance guarantees. Considering the uncertainty in estimating the multinomial logit choice model (MNL), Cheung and Simchi-Levi (2017) propose a Thompson-sampling-based policy to estimate the latent parameters by offering a personalized assortment, and they demonstrate its near optimality. We contribute to this literature by proposing a new, two-stage stochastic program framework to study the ad offer-set optimization problem. Moreover, we design a novel DWO policy that proves to be asymptotically optimal and generates values with lower variance than the FA benchmarks commonly adopted in the literature.

3. Model

3.1. Model Setup

The platform and its customers. We consider an e-commerce platform such as Amazon or Facebook Marketplace, which matches its user traffic with both organic product recommendations and advertisements. The advertisements thereof are usually labeled as *Sponsored Products*, as shown in Figure 1. Our model, however, can be straightforwardly applied to the setting of product recommendation on an e-commerce platform. Throughout the planning horizon, there are T customer impressions (also called user viewers, UVs) arriving at the platform sequentially. So we say customer t arrives in time t. Without loss of generality, we assume T is deterministic and known to the platform. Customers are segmented into m types based on their demographic information (e.g., age, gender, location) and behavior on the platform (e.g., average spending per year, product preferences, average time spent on the platform per year). We denote $\mathcal{M} := \{1, 2, ..., m\}$ as the set of all customer segments. For each customer t, her type $\tilde{j}(t)$ is *i.i.d.* and follows a discrete distribution on $\mathcal{M}$, with $\mathbb{P}(\tilde{j}(t) = j) = p^j > 0$ where $j \in \mathcal{M}$ and $\sum_{j\in\mathcal{M}} p^j = 1$.

Advertisements. At the beginning of the horizon, advertisers launch a set of ad campaigns, which we denote as $\mathcal{N} := \{1, 2, ..., n\}$. For each ad campaign $i \in \mathcal{N}$, its advertiser sets $B_i > 0$ as the total budget and $b_i > 0$ as the bid price. Specifically, B_i is the maximum advertising fee the advertiser will pay the platform throughout the ad campaign's life, and the budget will be depleted by b_i upon each *click* by a customer of the platform. That is, the platform adopts the CPC mechanism, which is commonly used in online advertising. The organic recommendation (i.e., the recommended items not labeled as "Sponsored" in Figure 1) is usually conducted by the recommendation department, a different team *not* responsible for advertising strategies. Furthermore, the ads are placed in some specific slots exclusively allocated to advertising. In this paper, we focus on the ad-assignment problem and, therefore, treat the organic recommendations as exogenous. Specifically, we use ad 0 to denote the representative organic recommendation, which is always displayed to each customer upon arriving at the platform. Hence, ad 0 may represent multiple products offered to the customer. We define $\bar{\mathcal{N}} = \mathcal{N} \cup \{0\}$. For ad 0 (i.e., the organic recommendation), the advertiser does not pay the platform to be displayed, so we set $b_0 = 0$ and $B_0 = +\infty$.

Offer-sets. Upon the arrival of customer t, the platform observes its type $j(t)$ and decides a (possibly randomized) set of ads/sponsored products displayed to this UV (which we call an *offer-set*), denoted by $\tilde{S}^{j(t)}(t) \in \mathfrak{S}^{j(t)} \subset 2^{\bar{\mathcal{N}}}$, where $\mathfrak{S}^j$ is the set of all feasible offer-sets to type j customers, and $2^{\bar{\mathcal{N}}}$ is the power set of $\bar{\mathcal{N}}$. We may impose additional structural constraints on $\mathfrak{S}^j$. Of particular importance is the cardinality constraint (i.e., the total number of ads displayed to customers cannot exceed K; e.g., Rusmevichientong et al. 2010, Wang 2012, Davis et al. 2013), which is prevalent in the online advertising setting where the platform cannot allocate more ads than the available ad-impression slots to a customer. We will demonstrate how to handle this cardinality constraint in our theoretical and numerical analyses below. Where it's not necessary to emphasize the realized customer type $j(t)$, we denote $\tilde{S}(t)$ as the offer-set displayed to customer t.

Click-throughs. For realized customer $j(t) \in \mathcal{M}$, if an offer-set $\tilde{S}^{j(t)}(t)$ is displayed (see, Figure 1), she may or may not click some ads in the set $\tilde{S}^{j(t)}(t)$. Since the organic recommendation is always included in the offer-set, we have $0 \in \tilde{S}^j(t)$ for each j. We denote $\tilde{y}_i^j(t)$ as the random variable that represents the total number of clicks received by ad i from a type j customer in time t. Therefore, $\tilde{y}_i^j(t) > 0$ only if $\tilde{j}(t) = j$, $i \in \tilde{S}^j(t)$, and customer t clicks ad i. In particular, $\tilde{y}_i^j(t) = 0$ if $\tilde{j}(t) \neq j$ or $i \notin \tilde{S}^j(t)$. We denote the click-through matrix in time t as $\tilde{\boldsymbol{y}}(t) := (\tilde{y}_i^j(t) : i \in \bar{\mathcal{N}}, j \in \mathcal{M})$, and its realization as $\boldsymbol{y}(t) := (y_i^j(t) : i \in \bar{\mathcal{N}}, j \in \mathcal{M})$. We do not specify any structure of the customer click-through behavior. Our model up to now is general and allows for multiple clicks of one or more ads from a customer. For each customer type j, each offer-set S, and each ad $i \in S$, we denote the expected click-throughs of ad i in an offer-set S from a type-j customer as $\phi_i^j(S) := \mathbb{E}_{\mathsf{D}_y}[\tilde{y}_i^j(t) | \tilde{j}(t) = j, \tilde{S}(t) = S]$, where D_y is the click-through distribution.

We assume customers exhibit (conditionally) independent and stationary click-through behaviors. Specifically, conditioned on the realized offer-sets $\{S(\tau) : 1 \leq \tau \leq T\}$, the click-throughs, $\tilde{\boldsymbol{y}}(t)$'s, are independent across time t. Furthermore, conditioned on the same offer-set, $\tilde{\boldsymbol{y}}(t)$'s are identically distributed with respect to the time t, i.e., for any set of click-through outcomes $\mathcal{Y}$, any realized customer type j, any offer-set S, and any $t \neq \tau$, $\mathbb{P}[\tilde{\boldsymbol{y}}(t) \in \mathcal{Y} | \tilde{j}(t) = j, \tilde{S}(t) = S] = \mathbb{P}[\tilde{\boldsymbol{y}}(\tau) \in \mathcal{Y} | \tilde{j}(\tau) = j, \tilde{S}(\tau) = S]$. By stationarity, the function $\phi_i^j(\cdot)$ is independent of time t. If $i \notin S$, $\phi_i^j(S) = 0$ by definition. Without loss of generality, we assume the expected number of click-throughs is finite, i.e., $\phi_i^j(S) < +\infty$ for any $i \in S \subset \bar{N}$ and $j \in \mathcal{M}$, which is sufficient for our theoretical results.

Ad targeting and click-through requirement. From the advertisers' perspective, they target ads to the relevant customer segments based on their demographic information, past behavioral patterns, and potential interests (see, e.g., Choi et al. 2020). Specifically, each ad campaign $i \in \mathcal{N}$ is characterized by a set of customer segments $\mathcal{L}_i \subseteq \mathcal{M}$ targeted by this ad. If the realized type of customer t, $j(t)$, is not targeted by ad i (i.e., $j(t) \notin \mathcal{L}_i$), the platform should not display ad i to this customer (i.e., $i \notin \tilde{S}(t)$). For the organic recommendation, the platform generally does not impose any constraints on the customer segments it may be displayed to, i.e., $\mathcal{L}_0 = \mathcal{M}$.

Furthermore, consistent with advertising practices (e.g., Microsoft 2020), the advertiser may require that ad i receives at least $\eta_i^{\mathcal{C}}$ click-throughs throughout the planning horizon for a set of customer types $\mathcal{C} \subseteq \mathcal{L}_i$. Mathematically, the targeting and click-through requirement of ad i is formalized as $\sum_{t=1}^{T} \sum_{j \in \mathcal{C}} \tilde{y}_i^j(t) \geq \eta_i^{\mathcal{C}}$ for any $i \in \mathcal{N}$ and $\mathcal{C} \subseteq \mathcal{L}_i$. For example, the advertiser may require the platform to target its ads for diapers to new parents.We use $\mathfrak{K}_i \subset 2^{\mathcal{L}_i}$ to denote the set of all customer segment subsets $\mathcal{C}$ on which the advertiser sets a positive click-through requirement $\eta_i^{\mathcal{C}} > 0$. In practice, $\mathfrak{K}_i$ often only contains either $\mathcal{L}_i$ itself (i.e. the requirement of total click-throughs) or some nonoverlapping subsets of $\mathcal{L}_i$. Moreover, large-scale online advertising platforms (e.g., Facebook and Google) have recently tightened their controls to prevent their advertisers from excluding some user segments in their target in order to reduce lawsuits and regulatory probes on discrimination. Thus, in most cases, the total number of click-through requirement constraints $|\mathfrak{K}_i|$ is at most *linear* (instead of exponential) in the total number of customer segments m, making our model and solution approach scalable. In many scenarios, the advertising contract specifies the minimal click-through requirement. For instance, Microsoft (as an advertiser) requires its partners (i.e., the advertising platforms where Microsoft runs its advertising campaigns) to earn at least 250 click-throughs during one ad campaign to be qualified to receive the support through its Co-op.

Advertising value and fairness. The total value of online advertising generated throughout the planning horizon depends on matching the n ads with T customers. Specifically, each click of ad i by a type-j customer generates a value of r_i^j, which is allowed to be both ad- and customer-type-specific. The interpretation of r_i^j, which can be quite general, includes the following scenarios as special cases. For the case where the value is the total advertising revenue of the platform, $r_i^j = b_i$ for each $i \in \mathcal{N}$. For the case where the value is the total advertising return of the advertisers (e.g., Hao et al. 2020), r_i^j is interpreted as the value of one click-through for ad i by a type-j customer to its advertiser. In particular, r_0^j is the value per click for an organic recommendation by a type-j customer. For example, r_0^j can be the average commission the seller pays for one click-through of the organic recommendation. For an e-commerce platform, r_0^j is in general one order of magnitude lower than r_i^j ($i \neq 0$). Therefore, the total value of online advertising is given by:

$$\sum_{t=1}^{T} \sum_{i \in \bar{\mathcal{N}}} \sum_{j \in \mathcal{L}_i} r_i^j \tilde{y}_i^j(t) \tag{1}$$

A salient feature of our model is that, in addition to the total advertising value, the platform may also concern the *fairness* of the system. For example, the recent advocacy on machine learning/algorithmic fairness postulates that customers of minority types should have sufficient click-throughs/conversions in a recommender/advertising system; otherwise, their needs cannot be well taken care of due to data sparsity (see, e.g., Lambrecht and Tucker 2019). In a similar vein, advertisers generally prefer receiving impressions that are evenly spread across their targeted customer types (e.g., Lejeune and Turner 2019). To account for such algorithmic fairness, we introduce a general fairness metric $F(\cdot) : \mathbb{R}^{(n+1)m} \mapsto \mathbb{R}$, which is a concave function of the per-customer click-through matrix: $\bar{\boldsymbol{y}} := (\bar{y}_i^j : i \in \bar{\mathcal{N}}, j \in \mathcal{M})$, where $\bar{y}_i^j := \frac{1}{T} \sum_{t=1}^{T} \tilde{y}_i^j(t)$ is the per-customer click-through of ad i by type-j customers. Because there is no universally accepted quantitative definition of fairness and equality, the fairness metric $F(\cdot)$ may one of many forms in different contexts, which we describe in detail below in Section 3.3. Therefore, we measure the fairness in a *per-customer* sense and evaluate the *per-customer* FAV from advertising as:

$$\frac{1}{T} \sum_{t=1}^{T} \sum_{i \in \bar{\mathcal{N}}} \sum_{j \in \mathcal{L}_i} r_i^j \tilde{y}_i^j(t) + \lambda F(\bar{\boldsymbol{y}}),$$

where $\lambda \geq 0$ parameterizes the trade-off between efficiency and fairness. The smaller (resp. larger) the λ, the higher weight the platform puts on efficiency (resp. fairness). In the extreme case where $\lambda \to 0$ (resp. $\lambda \to +\infty$), the system is purely efficiency-drive (resp. fairness-driven).

3.2. Stochastic Program Formulation

We consider the (randomized) non-anticipative policies $\tilde{\Pi}$. First, we define the realized history until time t as $\mathcal{H}_{t-1} := \{(j(\tau), S(\tau), \boldsymbol{y}(\tau)) : 1 \leq \tau \leq t-1\}$. By convention, $\mathcal{H}_0 = \emptyset$. For any policy $\tilde{\pi} \in \tilde{\Pi}$, we denote its implementation in time t as $\tilde{\pi}_t$. Essentially, $\tilde{\pi}_t$ maps the realized customer type $j(t)$ and the realized history $\mathcal{H}_{t-1}$ to a distribution on all the feasible offer-sets to a type-$j(t)$ customer, $\mathfrak{S}^{j(t)}$, i.e., $\tilde{S}^{j(t)}(t) = \tilde{\pi}_t(j(t), \mathcal{H}_{t-1})$. The deterministic non-anticipative policies, Π, are special cases of $\tilde{\Pi}$, which map $(j(t), \mathcal{H}_{t-1})$ to a deterministic offer-set in $\mathfrak{S}^{j(t)}$, i.e., $S^{j(t)}(t) = \pi_t(j(t), \mathcal{H}_{t-1})$ in time t under $\pi \in \Pi$. Sometimes, it is useful to spell out the dependence of the click-through outcomes in time t, $\tilde{\boldsymbol{y}}(t)$, on the history $\mathcal{H}_{t-1}$ and the policy $\tilde{\pi}_t$. We use $\tilde{y}_i^j(t|\tilde{\pi}_t, \mathcal{H}_{t-1})$ as the number of click-throughs for ad i by a type-j customer in time t, given that the history is $\mathcal{H}_{t-1}$ and the offer-set displayed to a type-$\tilde{j}(t)$ customer is $\tilde{S}^{\tilde{j}(t)}(t) = \tilde{\pi}_t(\tilde{j}(t), \mathcal{H}_{t-1})$. Likewise, we define $\bar{y}_i^j(\tilde{\pi}) := \frac{1}{T}\sum_{t=1}^{T} \tilde{y}_i^j(t|\tilde{\pi}_t, \mathcal{H}_{t-1})$ as the per-customer click-throughs of ad i by type j customers in the entire horizon under policy $\tilde{\pi}$. We denote $\bar{\boldsymbol{y}}(\tilde{\pi}) := (\bar{y}_i^j(\tilde{\pi}) : i \in \bar{\mathcal{N}}, j \in \mathcal{M})$ as the per-customer click-through matrix under policy $\tilde{\pi}$.

Of particular importance are the stationary history-independent policies, the set of which we denote as $\tilde{\Pi}_{\mathsf{s}} \subset \tilde{\Pi}$. Specifically, if $\tilde{\pi} \in \tilde{\Pi}_{\mathsf{s}}$, the offer-set displayed in time t, $\tilde{S}_t = \tilde{\pi}_t(j(t), \mathcal{H}_{t-1})$ is independent of (i) time t and (ii) the previous history $\mathcal{H}_{t-1}$ conditioned on the realized customer type $j(t)$. Hence, for $\tilde{\pi} \in \tilde{\Pi}_{\mathsf{s}}$, we can drop the time index t and history $\mathcal{H}_{t-1}$ to write $\tilde{\pi}_t(j(t), \mathcal{H}_{t-1})$ as $\tilde{\pi}(j(t))$ and $\tilde{y}_i^j(t|\tilde{\pi}_t, \mathcal{H}_{t-1})$ as $\tilde{y}_i^j(t|\tilde{\pi})$. Stationarity also implies that the distribution of $\tilde{S}^j(t)$ is stationary with respect to time t for each customer type j.

We are now ready to formulate the platform's ad-assignment problem as a multiperiod stochastic program.[6] Specifically, the platform seeks to optimize the total expected FAV of online advertising throughout the planning horizon:

$$
\begin{aligned}
&\max_{\tilde{\pi} \in \tilde{\Pi}} \mathbb{E}_{\tilde{\pi}, \tilde{\boldsymbol{j}}, \mathsf{D}_y} \left[\frac{1}{T} \sum_{t=1}^{T} \sum_{i \in \bar{\mathcal{N}}} \sum_{j \in \mathcal{L}_i} r_i^j \tilde{y}_i^j(t|\tilde{\pi}_t, \mathcal{H}_{t-1}) + \lambda F(\bar{\boldsymbol{y}}(\tilde{\pi})) \right] \\
&\text{s.t. } \frac{1}{T} \sum_{t=1}^{T} \sum_{j \in \mathcal{L}_i} b_i \tilde{y}_i^j(t|\tilde{\pi}_t, \mathcal{H}_{t-1}) \le \frac{B_i}{T}, \ \textit{almost surely} \text{ for each } i \in \mathcal{N}, \\
&\qquad \frac{1}{T} \sum_{t=1}^{T} \sum_{j \in \mathcal{C}} \tilde{y}_i^j(t|\tilde{\pi}_t, \mathcal{H}_{t-1}) \ge \frac{\eta_i^{\mathcal{C}}}{T}, \ \textit{almost surely} \text{ for each } i \in \bar{\mathcal{N}} \text{ and } \mathcal{C} \in \mathfrak{K}_i,
\end{aligned}
\tag{$\mathcal{OP}$}
$$

where the first term in the objective is the total per-customer value from advertising, which we call the *efficiency* of policy $\tilde{\pi}$ denoted by $\mathcal{E}(\tilde{\pi}) := \mathbb{E}_{\tilde{\pi}, \tilde{\boldsymbol{j}}, \mathsf{D}_y} \left[\frac{1}{T} \sum_{t=1}^{T} \sum_{i \in \bar{\mathcal{N}}} \sum_{j \in \mathcal{L}_i} r_i^j \tilde{y}_i^j(t|\tilde{\pi}_t, \mathcal{H}_{t-1}) \right]$, and the second term is the *fairness* of policy $\tilde{\pi}$ denoted by $\lambda \mathcal{F}(\tilde{\pi}) := \lambda \mathbb{E}_{\tilde{\pi}, \tilde{\boldsymbol{j}}, \mathsf{D}_y} [F(\bar{\boldsymbol{y}}(\tilde{\pi}))]$. Hence, the total FAV under policy $\tilde{\pi}$ is given by $\mathcal{V}(\tilde{\pi}) := \mathcal{E}(\tilde{\pi}) + \lambda \mathcal{F}(\tilde{\pi})$. We also remark that the first constraint of ($\mathcal{OP}$) refers to the budget constraint of each ad and the second refers to the click-through requirement for each ad with respect to different sets of customer types. We denote the optimal FAV of ($\mathcal{OP}$) as $\mathcal{V}^* = \limsup_{\tilde{\pi} \in \tilde{\Pi}} \mathcal{V}(\tilde{\pi})$ and the optimal policy (if exists) as $\tilde{\pi}^* = \arg\max_{\tilde{\pi} \in \tilde{\Pi}} \mathcal{V}(\tilde{\pi})$.

Roadmap to solve ($\mathcal{OP}$). For the rest of this paper, our goal is to design a policy $\tilde{\pi}$ that achieves the optimal FAV, $\mathcal{V}^*$, while satisfying the budget- and click-through-requirement constraints. As detailed below, our approach to solve ($\mathcal{OP}$) can be decomposed into the following two stages:

- **First-stage click-through target optimization,** where we solve a (deterministic but *nonequivalent*) convex program to obtain the optimal click-through target of the ads and customer types (Section 4).

[6] The derivation of our solution algorithm is based on the stochastic program ($\mathcal{OP}$) directly. Alternatively, the ad-assignment problem can be formulated as an equivalent dynamic program (DP), which is prohibitive. We refer interested readers to Appendix B for details.

- **Second-stage offer-set optimization,** where we adaptively decide the offer-set displayed to each customer based on how far away the click-throughs are from the optimal targets obtained in the first-stage (Section 5.1).

Altogether, we will demonstrate that our approach is asymptotically optimal (Section 5.2) and enjoys impressive performance in the nonasymptotic regime compared with the commonly adopted benchmarks in the existing literature (Section 6).

3.3. Metrics of Fairness

In this subsection, we describe a few commonly adopted metrics of fairness $F(\cdot)$, all of which can be coherently incorporated into our framework.

Max-min fairness. The recent trend of machine-learning fairness has promoted that minority customers should have sufficient click-throughs in a recommender/advertising system; otherwise, their needs cannot be well taken care of due to data scarcity. A natural choice to accommodate such fairness concern is the max-min fairness metric, which has been extensively studied in the literature of economics (e.g., Young and Isaac 1995), computer science (e.g., Kumar and Kleinberg 2000), and operations research (e.g., Bertsimas et al. 2012). Specifically, we define function $F(\cdot)$ as follows:

$$F(\bar{\boldsymbol{y}}) = \min_{j \in \mathcal{M}} \left\{ \sum_{i \in \mathcal{N}} \bar{y}_i^j \right\} \tag{2}$$

Max-min fairness drives the platform to maximize the minimum per-customer click-throughs from all customer types, ensuring that no customer type has too few click-throughs. We also note that max-min fairness can also be evaluated with respect to advertisers, i.e., $F(\bar{\boldsymbol{y}}) = \min_{i \in \mathcal{N}} \left\{ \sum_{j \in \mathcal{M}} \bar{y}_i^j \right\}$, so as to ensure that no advertiser receives too few click-throughs. One may also generalize max-min fairness to the $\alpha-$fairness metric (see, e.g., Bertsimas et al. 2012), i.e., $F(\bar{\boldsymbol{y}}) = \sum_{j \in \mathcal{M}} \frac{1}{1-\alpha} \left(\sum_{i \in \bar{\mathcal{N}}} \bar{y}_i^j \right)^{1-\alpha}$ if $\alpha \neq 1$ and $F(\bar{\boldsymbol{y}}) = \sum_{j \in \mathcal{M}} \log \left(\sum_{i \in \bar{\mathcal{N}}} \bar{y}_i^j \right)$ if $\alpha = 1$, which is reduced to the max-min fairness metric if we take $\alpha \to +\infty$.

Gini mean difference fairness. Advertisers generally prefer receiving impressions/click-throughs/conversions that are evenly spread across their targeted customer types (e.g., Lejeune and Turner 2019). One way to capture such preference is through Gini mean difference (GMD) fairness. The Gini coefficient/index has long been a canonical measure of income inequality in economics (e.g., Atkinson 1975), and it has recently been studied in the advertising literature to maximize the spreading of impressions across targeted UV types (e.g., Lejeune and Turner 2019). Following Lejeune and Turner (2019), given the average click-through matrix of all the ad/customer-type pairs, $\bar{\boldsymbol{y}}$, we define the GMD fairness metric as follows:

$$F(\bar{\boldsymbol{y}}) = -\sum_{i \in \bar{\mathcal{N}}} \frac{1}{\sum_{j \in \mathcal{L}_i} p^j} \sum_{j, j' \in \mathcal{L}_i} |p^{j'} \bar{y}_i^j - p^j \bar{y}_i^{j'}| \tag{3}$$

It is clear from (3) that the GMD fairness metric prompts the platform to induce click-throughs from each targeted customer type j proportional to its traffic p^j. For completeness, we provide a detailed derivation of the GMD fairness metric in Appendix C.

4. Reformulation and Feasibility Conditions

In this section, we propose a novel reformulation of the original intractable ad-allocation program $(\mathcal{OP})$ to maximize the expected FAV as a much simpler two-stage convex optimization. The core of our reformulation is to introduce the auxiliary click-through targets of the ads by different customer types, then design a novel, online, debt-based ad-allocation algorithm to achieve these targets. We also emphasize that although the reformulated two-stage convex optimization is *not* necessarily equivalent to $(\mathcal{OP})$, our proposed algorithms are indeed *provably optimal* for the original problem in the asymptotic regime where the problem size scales up to infinity.

4.1. Problem Reformulation With Click-Through Targets

The standard approach to solve the original dynamic ad-allocation problem $(\mathcal{OP})$ is to formulate it as a DP and obtain the optimal policy in each time t, $\tilde{\pi}_t^*$, by backward induction (see, e.g., Liu and Van Ryzin 2008, Bernstein et al. 2015). As we illustrate in Appendix B, the DP model that characterizes the optimal policy suffers from the curse of dimensionality and is, therefore, intractable from both analytical and computational perspectives. The other commonly adopted approach in the literature is to consider a fluid approximation (FA) of the problem and solve the FA problem by linear programming (Choice-based Linear Programming (CDLP), see, e.g., Liu and Van Ryzin 2008, in our case, the FA is a convex program) and linear-programming-resolving (LP-resolving) heuristics (LP-resolving, see, e.g., Jasin and Kumar 2012, Bumpensanti and Wang 2020). The FA-based formulation of $(\mathcal{OP})$ is provided by Eq. $\mathcal{FA}(\gamma)$ in Section 5.2 as an auxiliary problem to demonstrate the optimality of our proposed algorithm. With the cardinality constraint on the feasible offer-sets, one difficulty using (the analogs of) CDLP or LP-resolving is that the number of variables (i.e., the probability of each offer-set for all customer types) quickly explodes as the number of products increases, even when the choice model is as simple as the independent or MNL model.

To tackle the aforementioned challenges of standard approaches, we develop a novel reformulation of $(\mathcal{OP})$ that transforms the original problem as a two-stage convex optimization by introducing click-through targets associated with each ad-customer pair as auxiliary decision variables. Such reformulation also proves useful to design our asymptotically optimal ad-allocation algorithm. Specifically, we define $\boldsymbol{\alpha} := (\alpha_i^j, i \in \bar{\mathcal{N}}, j \in \mathcal{M}) \in \mathbb{R}_+^{(n+1)m}$, where α_i^j refers to the (virtual) target for the per-customer number of click-throughs for ad i by type j customers. So the platform operationalizes its ad-allocation algorithm such that the total number of click-throughs for ad i by type j customers exceeds $T\alpha_i^j$. We define the concave per-customer FAV associated with click-through target vector $\boldsymbol{\alpha}$ as

$$\mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}) := \sum_{i\in\bar{\mathcal{N}}}\sum_{j\in\mathcal{L}_i} r_i^j \alpha_i^j + \lambda F(\boldsymbol{\alpha}) \tag{4}$$

where the first term captures efficiency, the second captures fairness with respect to the click-through target vector $\boldsymbol{\alpha}$, and the subscript "CT" stands for "click-through target." We transform $(\mathcal{OP})$ into the following (nonequivalent) optimization problem:

$$\begin{aligned} \max_{\tilde{\pi}\in\tilde{\Pi},\boldsymbol{\alpha}\geq\mathbf{0}} \; & \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}) \\ \text{s.t. } & \frac{1}{T}\sum_{t=1}^{T} y_i^j(t|\tilde{\pi}_t,\mathcal{H}_{t-1}) \geq \alpha_i^j, \; \textit{almost surely} \text{ for each } i\in\bar{\mathcal{N}}, j\in\mathcal{M}, \\ & b_i\sum_{j\in\mathcal{L}_i}\alpha_i^j \leq \frac{B_i}{T}, \; \text{for each } i\in\mathcal{N}, \\ & \sum_{j\in\mathcal{C}}\alpha_i^j \geq \frac{\eta_i^{\mathcal{C}}}{T}, \quad \text{for each } i\in\bar{\mathcal{N}} \text{ and } \mathcal{C}\in\mathfrak{K}_i \end{aligned} \tag{5}$$

Comparing (5) with $(\mathcal{OP})$ reveals that our reformulation relaxes the sample-path-based objective function and constraints in the original problem with their counterparts characterized by the per-customer click-through target vector $\boldsymbol{\alpha}$. To ensure that the reformulation is close enough to the original problem and that the click-through targets are achievable, we introduce an additional sample-path constraint that the empirical click-throughs per customer should meet the click-through targets, as specified by the first constraint of (5). We emphasize that (5) is *not* equivalent to the original problem $(\mathcal{OP})$ in general. However, we show in Section 5 that in the asymptotic regime where the problem size scales up to infinity, there is an algorithm based on the solution to (5) achieving the optimal FAV of $(\mathcal{OP})$. In this sense, our reformulation is *asymptotically equivalent* to the original dynamic ad-allocation problem.

It is still challenging to characterize when the sample-path click-through target constraint in (5), i.e., $\frac{1}{T}\sum_{t=1}^{T} y_i^j(t|\tilde{\pi}_t,\mathcal{H}_{t-1}) \geq \alpha_i^j$, can be satisfied. Therefore, we further relax (5) by replacing this constraint with one for the expected number of click-throughs under stationary history-independent policies $\tilde{\Pi}_{\mathsf{s}}$. Specifically, we replace the first constraint of (5) with

$$\mathbb{E}_{\tilde{\pi},\tilde{j}(t),\mathsf{D}_y}[\tilde{y}_i^j(t|\tilde{\pi})] \geq \alpha_i^j, \quad \text{for each } i\in\bar{\mathcal{N}}, j\in\mathcal{M} \tag{6}$$

where $\tilde{\pi}\in\tilde{\Pi}_{\mathsf{s}}$. One should note that the expected click-through target constraint (6) is independent of time t. Hence, we reformulate the ad-allocation problem by further relaxing the click-through target constraint as follows:

$$\begin{aligned}
\max_{\tilde{\pi}\in\tilde{\Pi}_{\mathsf{s}},\boldsymbol{\alpha}\geq\mathbf{0}} \; & \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}) \\
\text{s.t. } & \mathbb{E}_{\tilde{\pi},\tilde{j}(t),\mathsf{D}_y}[\tilde{y}_i^j(t|\tilde{\pi})] \geq \alpha_i^j, \quad \text{for each } i\in\bar{\mathcal{N}}, j\in\mathcal{M}, \\
& b_i \sum_{j\in\mathcal{L}_i} \alpha_i^j \leq \frac{B_i}{T}, \text{ for each } i\in\mathcal{N}, \\
& \sum_{j\in\mathcal{C}} \alpha_i^j \geq \frac{\eta_i^{\mathcal{C}}}{T}, \quad \text{for each } i\in\bar{\mathcal{N}} \text{ and } \mathcal{C}\in\mathfrak{K}_i
\end{aligned} \tag{$2\mathcal{SSP}$}$$

It is clear from (2$\mathcal{SSP}$) that the original problem is reformulated as a *two-stage stochastic program.* In the first stage, the platform selects the click-through targets $\boldsymbol{\alpha}$ to maximize a variant of the FAV, $\mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha})$; in the second stage, it selects a stationary history-independent policy $\tilde{\pi}$ to meet the first-stage click-through targets (6).

We call a click-through target vector $\boldsymbol{\alpha}$ a *single-period feasible* if there exists a stationary history-independent policy $\tilde{\pi}\in\tilde{\Pi}_{\mathsf{s}}$ such that (6) holds. The single-period feasibility condition for a click-through target vector $\boldsymbol{\alpha}$ is central to the design and analysis of our algorithm to solving both the two-stage formulation (2$\mathcal{SSP}$) and, eventually, the original problem ($\mathcal{OP}$). The rest of this section will be devoted to characterizing the necessary and sufficient condition for an $\boldsymbol{\alpha}$ to be single-period feasible.

Sometimes it is more convenient to rewrite this expected click-through target condition (6) as a periodic-review infinite-horizon sample average-feasibility condition, which will prove useful to establish the optimal dynamic ad-allocation policy, i.e., to find a (randomized) non-anticipative policy $\tilde{\pi}$, such that

$$\liminf_{T\uparrow+\infty} \frac{1}{T}\sum_{t=1}^{T} \tilde{y}_i^j(t|\tilde{\pi}_t,\mathcal{H}_{t-1}) \geq \alpha_i^j, \quad \text{for each } i\in\bar{\mathcal{N}}, j\in\mathcal{M} \tag{7}$$

Note that a similar periodic-review reformulation of service-level constraints has also been adopted in the literature on resource allocation and inventory pooling (e.g. Zhong et al. 2017, Jiang et al. 2019).

4.2. Necessary and Sufficient Condition for Single-Period Feasibility

To obtain the optimal click-through targets that solve (2$\mathcal{SSP}$), we first characterize the necessary and sufficient condition under which the first-stage click-through target vector $\boldsymbol{\alpha}$ is single-period feasible, i.e., (6) holds. We consider the following formulation with a constant objective function:

$$\begin{aligned}
\max_{\tilde{\pi}\in\tilde{\Pi}_{\mathsf{s}}} \; & 0 \\
\text{s.t. } & \mathbb{E}_{\tilde{\pi},\tilde{j}(t),\mathsf{D}_y}[\tilde{y}_i^j(t|\tilde{\pi})] \geq \alpha_i^j, \quad \text{for each } i\in\bar{\mathcal{N}}, j\in\mathcal{M}
\end{aligned} \tag{8}$$

Note that, due to stationarity, (8) is regardless of time t. We now characterize when the stochastic program (8) has a feasible solution. We first reformulate (8) as a linear program (LP). Note that the set of deterministic stationary history-independent policies Π_{s} are all the mappings that take a type-j customer to an offer-set in $\mathfrak{S}^j$, which is finite with cardinality $|\Pi_{\mathsf{s}}| = \prod_{j\in\mathcal{M}} |\mathfrak{S}^j|$. Hence, a randomized policy $\tilde{\pi} \in \tilde{\Pi}_{\mathsf{s}}$ is defined by a probability measure $\mu(\cdot)$ on the finite set Π_{s}, which is essentially a probability simplex in the space $\mathbb{R}^{|\Pi_{\mathsf{s}}|}$.

Under a deterministic, stationary, history-independent policy $\pi \in \Pi_{\mathsf{s}}$, if a type-j customer arrives, the platform displays an offer-set $S^j = \pi(j)$ (due to stationarity, we drop the time index t). Thus, the average per-customer number of click-throughs for ad i by type-j customers is given by

$$p^j \phi_i^j(\pi(j))$$

Therefore, (8) can be reformulated as the following LP, the solution to which we denote as $\mu^*(\cdot)$:

$$\begin{aligned} \max_{\mu(\cdot)}\ & 0 \\ \text{s.t. } & \sum_{\pi\in\Pi_{\mathsf{s}}} \mu(\pi) p^j \phi_i^j(\pi(j)) \geq \alpha_i^j, \text{ for each } i \in \bar{\mathcal{N}} \text{ and } j \in \mathcal{M} \\ & \sum_{\pi\in\Pi_{\mathsf{s}}} \mu(\pi) = 1 \\ & \mu(\pi) \geq 0 \text{ for all } \pi \in \Pi_{\mathsf{s}} \end{aligned} \tag{9}$$

Taking the dual of the LP (9), we obtain that

$$\begin{aligned} \min_{\theta_0, \theta_i^j}\ & \{\theta_0 - \sum_{i\in\bar{\mathcal{N}}, j\in\mathcal{M}} \alpha_i^j \theta_i^j\} \\ \text{s.t. } & \sum_{i\in\bar{\mathcal{N}}, j\in\mathcal{M}} p^j \phi_i^j(\pi(j))\theta_i^j - \theta_0 \leq 0, \text{ for all } \pi \in \Pi_{\mathsf{s}} \\ & \theta_i^j \geq 0 \text{ for all } i \in \bar{\mathcal{N}} \text{ and } j \in \mathcal{M} \end{aligned} \tag{10}$$

Note that, in (10), $\theta_j^i \geq 0$ is the shadow price for the click-through target of ad i from customer segment j, $\sum_{\pi\in\Pi_{\mathsf{s}}} \mu(\pi) p^j \phi_i^j(\pi(j)) \geq \alpha_i^j$, whereas θ_0 is the dual-variable for the normalization constraint $\sum_{\pi\in\Pi_{\mathsf{s}}} \mu(\pi) = 1$. We also define $\boldsymbol{\theta} := (\theta_i^j : i \in \bar{\mathcal{N}}, j \in \mathcal{M})$.

We note that the objective function of the primal formulation (9) is a constant 0, and there exists a feasible solution $\theta_0 = 0$ and $\theta_i^j = 0$ (for all i and j) to the dual formulation (10) with objective value equal to 0. On one hand, if (9) has a feasible solution, the dual formulation (10) will have the optimal objective value greater than or equal to 0 by weak duality. On the other hand, if the optimal objective value of (10) is greater than or equal to 0, it should be exactly equal to 0, since the feasible solution with $\theta_0 = 0$ and $\theta_i^j = 0$ (for all i and j) generates a zero objective function value. In this case, strong duality implies that (9) will have a feasible solution. Therefore, putting the argument above together, (9) has a feasible solution if and only if the optimal objective function value of (10) is non-negative.

Moreover, θ_0 is non-negative by (10). Thus, the minimal objective function value of (10) can be obtained at the smallest feasible θ_0, which is $\max_{\pi\in\Pi_s} \sum_{i\in\bar{\mathcal{N}},j\in\mathcal{M}} p^j\phi_i^j(\pi(j))\theta_i^j$ based on the first set of constraints in (10). Combining the aforementioned two observations, we have the feasibility of (9) is equivalent to

$$\min_{\theta_i^j\geq 0}\left\{\max_{\pi\in\Pi_s}\sum_{i\in\bar{\mathcal{N}},j\in\mathcal{M}} p^j\phi_i^j(\pi(j))\theta_i^j - \sum_{i\in\bar{\mathcal{N}},j\in\mathcal{M}}\alpha_i^j\theta_i^j\right\}\geq 0$$

which is equivalent to

$$\max_{\pi\in\Pi_s}\sum_{i\in\bar{\mathcal{N}},j\in\mathcal{M}} p^j\phi_i^j(\pi(j))\theta_i^j \geq \sum_{i\in\bar{\mathcal{N}},j\in\mathcal{M}}\alpha_i^j\theta_i^j \text{ for all } \theta_i^j\geq 0,\ i\in\bar{\mathcal{N}}, j\in\mathcal{M} \tag{11}$$

The following theorem summarizes the above argument, and it establishes the necessary and sufficient condition for the click-through target vector $\boldsymbol{\alpha}$.

Theorem 1. (Necessary and Sufficient Condition) *A click-through target vector $\boldsymbol{\alpha}$ is single-period feasible, i.e., there exists a stationary history-independent policy $\tilde{\pi}\in\tilde{\Pi}_s$ such that (6) holds, if and only if (11) holds.*

Indeed, when $\boldsymbol{\alpha}$ satisfies (11), an optimal dual vector $\boldsymbol{\theta}^*$ that solves (10) helps characterize the set of *deterministic* stationary history-independent policies over which a primal policy $\mu^*(\cdot)$ (feasible to (9)) randomizes. Specifically, strong duality and the complementary slackness condition imply that for a deterministic policy $\pi\in\Pi_s$ to have a positive weight in a feasible primal policy $\mu^*(\cdot)$, i.e., $\mu^*(\pi)>0$, it must hold that the first constraint of the dual problem (10) is binding for π, i.e.,

$$\pi(j)\in\arg\max_{S\in\mathfrak{S}^j}\sum_{i\in S}p^j\theta_i^j\phi_i^j(S) = \arg\max_{S\in\mathfrak{S}^j}\sum_{i\in S}\theta_i^j\phi_i^j(S)$$

Note that the left-hand side of inequality (11) can be viewed as a personalized offer-set optimization problem. Specifically, for each customer type j, we seek to provide an offer-set S^{j*} that maximizes the total revenue from this customer type with his/her per-click revenue of ad i set to θ_i^j, i.e.,

$$S^{j*}(\boldsymbol{\theta}) = \arg\max_{S\in\mathfrak{S}^j}\sum_{i\in S}\theta_i^j\phi_i^j(S) \tag{12}$$

with ties broken arbitrarily so that $S^{j*}(\boldsymbol{\theta})$ is uniquely determined for a given $\boldsymbol{\theta}$. Given a vector $\boldsymbol{\theta}$, we denote the deterministic policy generated by (12) as $\pi_{\boldsymbol{\theta}}$ (hence, $\pi_{\boldsymbol{\theta}}(j)=S^{j*}(\boldsymbol{\theta})$). For an $\boldsymbol{\alpha}$ that satisfies (11), wisely generating a random vector $\tilde{\boldsymbol{\theta}}$ could produce a feasible randomized policy $\tilde{\pi}=\pi_{\tilde{\boldsymbol{\theta}}}$. We detail the procedure of generating $\tilde{\boldsymbol{\theta}}$ in Section 5.

Given the dual vector $\boldsymbol{\theta}$, we define $g(\boldsymbol{\theta}) := \sum_{i\in\bar{\mathcal{N}},j\in\mathcal{M}} p^j\theta_i^j\phi_i^j(S^{j*}(\boldsymbol{\theta}))$, which is the left-hand side of (11). Hence, we obtain an equivalent necessary and sufficient condition for the feasibility of click-through targets $\boldsymbol{\alpha}$:

$$\begin{aligned} & h(\boldsymbol{\alpha}) \geq 0 \\ \text{where } & h(\boldsymbol{\alpha}) := \min_{\boldsymbol{\theta}\geq\mathbf{0}} \left\{ g(\boldsymbol{\theta}) - \sum_{i\in\bar{\mathcal{N}},j\in\mathcal{M}} \alpha_i^j\theta_i^j : \theta_i^j \geq 0, \text{ for all } i\in\bar{\mathcal{N}}, j\in\mathcal{M} \right\} \end{aligned} \tag{13}$$

Because $g(\boldsymbol{\theta})$ is the maximum of a family of linear functions, it is jointly convex in $\boldsymbol{\theta}$. Therefore, checking the feasibility of the two-stage stochastic program (2$\mathcal{SSP}$) is reduced to minimizing a convex function $g(\boldsymbol{\theta}) - \sum_{i\in\bar{\mathcal{N}},j\in\mathcal{M}} \alpha_i^j\theta_i^j$ over the quadrant $\{\theta_i^j \geq 0 : i\in\bar{\mathcal{N}}, j\in\mathcal{M}\}$. Hence, as long as the personalized offer-set optimization problem (12) is tractable (e.g., the customer click-throughs follow independent, MNL, nested-MNL, or the generalized attraction models), one could numerically check the feasibility of the click-through targets $\boldsymbol{\alpha}$. By Eq. (13), $h(\boldsymbol{\alpha})$ is the minimum of a family of linear functions (in $\boldsymbol{\alpha}$), so it is jointly concave in $\boldsymbol{\alpha}$.

With the characterization of the necessary and sufficient condition (13) for the feasibility of click-through targets $\boldsymbol{\alpha}$ in the second stage, we are now ready to reformulate the two-stage stochastic program (2$\mathcal{SSP}$) as the following single-stage (*deterministic*) convex program to obtain the optimal target problem:

$$\begin{aligned} \max_{\boldsymbol{\alpha}\geq\mathbf{0}} \ & \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}) \\ \text{s.t. } & h(\boldsymbol{\alpha}) \geq 0, \\ & b_i \sum_{j\in\mathcal{L}_i} \alpha_i^j \leq \frac{B_i}{T}, \text{ for each } i\in\mathcal{N}, \\ & \sum_{j\in\mathcal{C}} \alpha_i^j \geq \frac{\eta_i^{\mathcal{C}}}{T}, \text{ for each } i\in\bar{\mathcal{N}} \text{ and } \mathcal{C}\in\mathfrak{K}_i \end{aligned} \tag{$\mathcal{OTP}$}$$

Of particular importance is a special case of ($\mathcal{OTP}$) where customer click-throughs follow the MNL choice model, i.e., for any $i\in S\in\mathfrak{S}^j$,

$$\phi_i^j(S) = \frac{v_i^j}{1+\sum_{i'\in S} v_{i'}^j} \tag{14}$$

where $v_i^j > 0$ is the attractiveness of ad i to type-j customers. We demonstrate in the following proposition that if customer click-throughs follow the MNL choice model (14) and the cardinality constraint for any offer-set displayed to a customer (i.e., $|\tilde{S}(t)| \leq K$ for some K), the optimal target problem ($\mathcal{OTP}$) can be simplified to a convex program with a few linear constraints.[7]

[7] In Appendix F, we consider that customer click-throughs follow the independent choice model (which is widely adopted in practice; see, e.g., Feldman et al. 2021) or the generalized attraction choice model (which is more general than MNL; see, e.g., Luce 2012, Gallego et al. 2015). We are able to show that the optimal target problem ($\mathcal{OTP}$) can also be simplified to tractable convex programs under independent and generalized attraction choice models.

Proposition 1. *If customer click-throughs follow the MNL choice model* (14) *and the size of an offer-set cannot exceed K, the first-stage convex program* $(\mathcal{OTP})$ *can be simplified to the following one:*

$$\begin{aligned}
&\max_{\boldsymbol{\alpha}\geq\mathbf{0}} \ \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}) \\
&s.t. \ \sum_{i'=1}^{n}\alpha_{i'}^{j} + \frac{(1+v_0^j)\alpha_i^j}{v_i^j} \leq p^j, \ \textit{for each } i\in\bar{\mathcal{N}}, j\in\mathcal{M}, \\
&\quad \sum_{i=1}^{n}\alpha_i^j + \frac{1+v_0^j}{K-1}\sum_{i=1}^{n}\frac{\alpha_i^j}{v_i^j} \leq p^j, \ \textit{for each } j \\
&\quad b_i\sum_{j\in\mathcal{L}_i}\alpha_i^j \leq \frac{B_i}{T}, \ \textit{for each } i\in\mathcal{N}, \\
&\quad \sum_{j\in\mathcal{C}}\alpha_i^j \geq \frac{\eta_i^{\mathcal{C}}}{T}, \quad \textit{for each } i\in\bar{\mathcal{N}} \textit{ and } \mathcal{C}\in\mathfrak{K}_i
\end{aligned} \tag{$\mathcal{OTP}-\mathcal{MNL}$}$$

Proposition 1 shows that the number of linear constraints for the convex program $(\mathcal{OTP}-\mathcal{MNL})$ is *linear* (instead of exponential) in m and n, which ensures its tractability.

Definition 1. We say that a click-through target vector $\boldsymbol{\alpha}\in\mathbb{R}_+^{(n+1)m}$ is *feasible* if it is a feasible solution to $(\mathcal{OTP})$.

By definition, if $\boldsymbol{\alpha}$ is *feasible*, then it is *single-period feasible*. Throughout this paper, we assume the *feasible* region is nonempty, so an optimal solution to $(\mathcal{OTP})$ exists, which we denote as $\boldsymbol{\alpha}^*$. Also, we denote $\mathcal{V}_{\mathsf{CT}}^* := \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}^*)$ as the optimal objective function value of $(\mathcal{OTP})$. Thus, $\boldsymbol{\alpha}^*$ is the "optimal" click-through target vector for our reformulated ad-allocation problem. According to Theorem 1, $h(\boldsymbol{\alpha})$ defined by (13) being non-negative provides a necessary and sufficient condition for the click-through targets, $\boldsymbol{\alpha}$, to be obtainable in the expected sense $\mathcal{V}_{\mathsf{CT}}^*$, which proves to be an upper bound of the optimal FAV for the original problem, $\mathcal{V}^*$ (see Theorem 3 below). The convex program formulation $(\mathcal{OTP})$, therefore, characterizes the optimal click-through target vector $\boldsymbol{\alpha}^*$ and the associated optimal (relaxed) per-customer FAV in the expected sense. However, two critical questions remain to be addressed:

- **Achieving $\boldsymbol{\alpha}^*$.** How should we display the offer-sets upon the arrival of each customer to achieve the optimal click-through goals $\boldsymbol{\alpha}^*$?
- **Optimality of achieving $\boldsymbol{\alpha}^*$.** Will the offer-set display strategy achieving $\boldsymbol{\alpha}^*$ suffice to obtain the true (nonrelaxed) optimal value of the original problem $(\mathcal{OP})$, i.e., $\mathcal{V}^*$?

The rest of this paper is devoted to addressing both questions.

5. Algorithms for Advertisement Allocation Optimization

In this section, we focus on devising a novel algorithm that addresses the ad-allocation problem $(\mathcal{OP})$ based on the solution to the click-through target optimization, $\boldsymbol{\alpha}^*$. Specifically, we propose an adaptive offer-set policy that meets the optimal click-through targets $\boldsymbol{\alpha}^*$, and we demonstrate that our proposed algorithm is asymptotically optimal as the problem size scales to infinity. If only a compromised solution can be obtained for the optimal target problem ($\mathcal{OTP}$), our proposed algorithms will achieve the same (asymptotic) optimality gap as that in the optimal target problem, suggesting the robustness of our approach. Our proposed algorithms have better scalability than DP- and FA- based approaches, and they can deplete the budget of each ad more smoothly throughout the horizon, which is highly desirable for the advertising business in practice.

5.1. Debt-Weighted Advertisement Allocation Policy

By our two-stage stochastic program (re)formulation of the ad-allocation problem, ($2\mathcal{SSP}$), once we solve the optimal click-through target vector $\boldsymbol{\alpha}^*$, the problem is reduced to devising a randomized offer-set algorithm to achieve $\boldsymbol{\alpha}^*$. To this end, one may solve the primal-dual problems (8) and (10) with $\boldsymbol{\alpha} = \boldsymbol{\alpha}^*$ to obtain a feasible randomized policy that achieves $\boldsymbol{\alpha}^*$. This approach, though intuitive, may be computationally prohibitive, because the primal LP (9) has $\mathcal{O}(m2^n)$ decision variables and $\mathcal{O}(mn)$ constraints. Therefore, we resort to a data-driven algorithm to generate the random dual vector $\tilde{\boldsymbol{\theta}}(t)$ upon the arrival of each customer t, based on which we adaptively customize the appropriate ad offer-set $S^{j(t)*}(\tilde{\boldsymbol{\theta}}(t))$. Algorithm 1 below presents our policy. We refer to the DWO policy (Algorithm 1) initialized with the click-through target vector $\boldsymbol{\alpha}$ as the DWO-$\boldsymbol{\alpha}$ policy, denoted by $\tilde{\pi}_{\mathsf{DWO}}(\boldsymbol{\alpha})$. Of particular importance is the DWO-$\boldsymbol{\alpha}^*$ policy, where the platform solves the optimal target problem ($\mathcal{OTP}$) *offline* to obtain the optimal click-through target vector $\boldsymbol{\alpha}^*$, then implements $\tilde{\pi}_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)$ *online* to adaptively display personalized offer-set to each customer. The main result of this section is that the DWO-$\boldsymbol{\alpha}^*$ policy is asymptotically optimal for the original ad-allocation problem ($\mathcal{OP}$).

Algorithm 1 DEBT-WEIGHTED OFFER-SET POLICY $\tilde{\pi}_{\mathsf{DWO}}(\boldsymbol{\alpha})$

Initialize: The click-through target vector $\boldsymbol{\alpha}$ and the initial debts $d_i^j(1) \leftarrow 0$ for all $i \in \bar{\mathcal{N}}$ and $j \in \mathcal{M}$.

For each customer $t \geq 1$**:**

1: Observe the customer-type $j(t)$.

2: Display the offer-set $S^{j(t)*}(\boldsymbol{d}(t))$ to customer t (see Eq. (12)), where $\boldsymbol{d}(t) = (d_i^j(t) : i \in \bar{\mathcal{N}}, j \in \mathcal{M})$ is the realized debt vector upon the arrival of customer t.

3: Observe the customer click-throughs $(y_i^j(t) : i \in \bar{\mathcal{N}}_s, j \in \mathcal{M})$. The advertising value, $\sum_i r_i^j y_i^{j(t)}(t)$, is collected. In the case where the budget for ad i is exhausted, i.e., $\left(\sum_j \sum_{\tau \leq t} y_i^j(\tau)\right) b_i \geq B_i$, any offer-set containing this ad is removed from $\mathfrak{S}^j$ for all j hereafter.

4: $d_i^j(t+1) \leftarrow d_i^j(t) + \alpha_i^j - y_i^j(t)$ for all $i \in \bar{\mathcal{N}}$ and $j \in \mathcal{M}$.

A few remarks are in order with respect to Algorithm 1. First, the DWO policy displays the offer-set to each customer based on the offer-set optimization problem (12). This standard personalized offer-set optimization problem is tractable so the offer-set $S^{j(t)*}(\boldsymbol{d}(t))$ can be efficiently obtained for a broad class of choice models: independent, MNL, nested MNL, and generalized attraction. Second, we call Algorithm 1 the DWO policy, because the offer-set optimization is weighted by the "debt" of each customer-advertisement pair for customers $\{1,2,...,t-1\}$. Note that $(t-1)\alpha_i^j$ is the total click-through target of ad i by type j customers until the start of time t, whereas $\sum_{\tau=1}^{t-1} y_i^j(\tau)$ is the total realized click-throughs by then. Therefore, $\left(d_i^j(t)\right)^+ = \max\left((t-1)\alpha_i^j - \sum_{\tau=1}^{t-1} y_i^j(\tau), 0\right)$ is the total "debt" owed by the platform to the click-through target associated with ad i and customer type j when deciding the offer-set displayed to customer t. For a feasible click-through target vector $\boldsymbol{\alpha}$, we can also view the debt process $\{\tilde{\boldsymbol{d}}(t) : t \geq 1\}$ as a data-driven adaptive way to generate the random dual vector $\tilde{\boldsymbol{\theta}}$, which prescribes a feasible randomized policy $\tilde{\pi} = \pi_{\tilde{\boldsymbol{\theta}}}$. Finally, note that the debts at the start of time t, $\tilde{\boldsymbol{d}}(t)$ only depend on $\mathcal{H}_{t-1}$ and are independent of any information revealed on or after time t. Therefore, the DWO policy is non-anticipative.

5.2. Asymptotic Analysis

In this subsection, we will establish that the DWO-$\boldsymbol{\alpha}^*$ policy can achieve the optimal FAV for the original ad-allocation problem ($\mathcal{OP}$) asymptotically. Before demonstrating the optimality of the DWO-$\boldsymbol{\alpha}^*$ policy, we first introduce the asymptotic regime where the problem size scales up to infinity. Specifically, we denote a family of ad-allocation problems with the budget for each ad i, $B_i(\gamma) := B_i\gamma$, the click-through requirement for ad i and customer-type set $\mathcal{C} \in \mathfrak{K}_i$, $\eta_i^{\mathcal{C}}(\gamma) = \eta_i^{\mathcal{C}}\gamma$, and the planning horizon length $T(\gamma) := T\gamma$, as $\mathcal{OP}(\gamma)$, where $\gamma > 0$ is a scaling parameter of problem size. Hence, the original problem ($\mathcal{OP}$) is equivalent to $\mathcal{OP}(1)$. For the problem $\mathcal{OP}(\gamma)$ and a policy $\tilde{\pi} \in \tilde{\Pi}$, we denote $\mathcal{E}(\tilde{\pi}|\gamma)$ as the expected efficiency, $\mathcal{F}(\tilde{\pi}|\gamma)$ as the expected fairness, and $\mathcal{V}(\tilde{\pi}|\gamma) = \mathcal{E}(\tilde{\pi}|\gamma) + \lambda\mathcal{F}(\tilde{\pi}|\gamma)$ as the expected FAV generated by $\tilde{\pi}$ in $\mathcal{OP}(\gamma)$. Furthermore, $\mathcal{V}^*(\gamma) := \max_{\tilde{\pi}\in\tilde{\Pi}} \mathcal{V}(\tilde{\pi}|\gamma)$ denotes the optimal expected FAV for $\mathcal{OP}(\gamma)$. Note that the market-size scaling factor γ does not affect the feasibility of a click-through target vector $\boldsymbol{\alpha}$, nor does it change the two-stage stochastic program reformulation ($2\mathcal{SSP}$) or the target problem reformulation ($\mathcal{OTP}$).

We first establish that, for any *feasible* click-through target vector $\boldsymbol{\alpha}$, the DWO-$\boldsymbol{\alpha}$ policy exactly achieves $\boldsymbol{\alpha}$ in $\mathcal{OP}(\gamma)$ as the problem size γ scales to infinity.

THEOREM 2. *If $\boldsymbol{\alpha}$ is feasible, i.e., all constraints of ($\mathcal{OTP}$) are satisfied, then we have:*

$$\lim_{\gamma\uparrow+\infty} \frac{1}{T(\gamma)} \sum_{t=1}^{T(\gamma)} \tilde{y}_i^j(t|\tilde{\pi}_{\mathsf{DWO}}(\boldsymbol{\alpha}), \mathcal{H}_{t-1}) = \alpha_i^j \text{ almost surely for all } i \in \bar{\mathcal{N}} \text{ and } j \in \mathcal{M}. \tag{15}$$

Theorem 2 is a central technical result of this paper, because it serves as an important stepping stone to prove the asymptotic optimality of the DWO-$\boldsymbol{\alpha}^*$ policy. Interestingly, as long as this policy is initiated with a *feasible* click-through target vector $\boldsymbol{\alpha}$, it will not only achieve click-through levels *at least as high as* these targets (i.e., Eq. (7)) but also *exactly approach* them (i.e., Eq. (15)). Adopting a coupling argument, the proof of Theorem 2 (see Appendix D for details) demonstrates that if the problem size γ scales up to infinity, the DWO-$\boldsymbol{\alpha}$ policy will *not* exhaust the budget of any ad and, thus, will secure the click-through targets $\boldsymbol{\alpha}$. Therefore, the click-through requirements $\{\eta_i^{\mathcal{C}}(\gamma) : i \in \mathcal{N}, \mathcal{C} \subset \mathcal{L}_i\}$ can be satisfied in the asymptotic regime as well.

Based on Theorem 2, one may conjecture that, if the click-through target vector $\boldsymbol{\alpha}$ is optimally chosen (i.e., as the solution to the optimal target problem ($\mathcal{OTP}$), $\boldsymbol{\alpha}^*$), the DWO-$\boldsymbol{\alpha}^*$ policy could achieve the optimal FAV for the original ad-allocation problem, $\mathcal{V}^*$. The main result of this section is that the following theorem validates this conjecture in the asymptotic regime.

THEOREM 3. *The* DWO-$\boldsymbol{\alpha}^*$ *policy is asymptotically optimal, i.e.,*

$$\lim_{\gamma \uparrow +\infty} \mathcal{V}(\tilde{\pi}_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma) = \lim_{\gamma \uparrow +\infty} \mathcal{V}^*(\gamma) = \mathcal{V}^*_{\mathsf{CT}} \tag{16}$$

Furthermore, the optimal objective function value of the first-stage click-through target optimization ($\mathcal{OTP}$) is an upper bound for the original problem ($\mathcal{OP}$) in the nonasymptotic regime, i.e., for any $\gamma > 0$*,*

$$\mathcal{V}^*_{\mathsf{CT}} \geq \mathcal{V}^*(\gamma) \tag{17}$$

Theorem 3 proves that the DWO-$\boldsymbol{\alpha}^*$ policy is asymptotically optimal when the ad budgets, the click-through requirements, and the time-horizon length all scale up to infinity at the same rate. In particular, the optimal expected FAV of $\mathcal{OP}(\gamma)$ is identical to the optimal FAV of the optimal target problem ($\mathcal{OTP}$) asymptotically, and the former is upper-bounded by the latter in the nonasymptotic regime. Such equivalence suggests that our reformulation in Section 4.1 is an asymptotically equivalent relaxation of the original problem. We highlight that our study provides the *optimality* guarantee of a *debt-weighted* policy in addition to its *feasibility/approchability*. In the existing literature, debt-weighted algorithms in the contexts of resource pooling, (e.g., Zhong et al. 2017, Jiang et al. 2019), and process flexibility (e.g., Lyu et al. 2019) have been validated to satisfy feasibility conditions of different formats. We prove that the DWO-$\boldsymbol{\alpha}^*$ policy which achieves the asymptotically optimal FAV is subject to the ad budget and click-through requirement constraints. Furthermore, to our best knowledge, we are also the first in the literature to study the dynamic assortment/offer-set optimization problem through the lens of a debt-weighted algorithm.

The proof of Theorem 3 also presents an interesting technical contribution of our paper. Leveraging the *exact approchability* result of Theorem 2, we establish the intrinsic connections and, therefore, the asymptotic equivalence of the original problem $\mathcal{OP}(\gamma)$, the optimal target problem $(\mathcal{OTP})$, and an auxiliary FA convex program, which is a generalization of the standard CDLP approach (Liu and Van Ryzin 2008). Establishing the connections and equivalence of these problems is not only essential to our proof, but also novel to the literature. The auxiliary FA problem with scale factor γ is defined as the convex program $\mathcal{FA}(\gamma)$:

$$
\begin{aligned}
\max_{\boldsymbol{z}} \ & \mathcal{V}_{\mathsf{FA}}(\boldsymbol{z}) := \sum_{i\in\bar{\mathcal{N}}, j\in\mathcal{M}, S\in\mathfrak{S}^j} r_i^j p^j \phi_i^j(S) z^j(S) + \lambda F(\boldsymbol{\zeta}) \\
\text{s.t.} \ & \sum_{j\in\mathcal{M}, S\in\mathfrak{S}^j} b_i p^j \phi_i^j(S) z^j(S) \le \frac{B_i(\gamma)}{T(\gamma)} \text{ for all } i\in\mathcal{N} \\
& \sum_{j\in\mathcal{C}, S\in\mathfrak{S}^j} p^j \phi_i^j(S) z^j(S) \ge \frac{\eta_i^{\mathcal{C}}(\gamma)}{T(\gamma)} \text{ for all } i\in\bar{\mathcal{N}} \text{ and } \mathcal{C}\in\mathfrak{K}_i \\
& \sum_{S\in\mathfrak{S}^j} z^j(S) \le 1 \text{ for all } j\in\mathcal{M} \\
& z^j(S) \ge 0 \text{ for all } j\in\mathcal{M}, S\in\mathfrak{S}^j \\
& \boldsymbol{\zeta}\in\mathbb{R}^{(n+1)m}, \text{ with } \zeta_i^j = \sum_{i\in S\in\mathfrak{S}^j} p^j z^j(S)\phi_i^j(S)
\end{aligned}
\qquad \mathcal{FA}(\gamma)
$$

It is self-evident from the formulation of $\mathcal{FA}(\gamma)$ that $z^j(S)$ is the probability of displaying offer-set S to a type j customer upon her arrival, whereas $\zeta_i^j = \sum_{i\in S\in\mathfrak{S}^j} p^j z^j(S)\phi_i^j(S)$ is the expected per-customer click-throughs of ad i by type-j customers. A vector $\boldsymbol{z} = (z^j(S): j\in\mathcal{M}, S\in\mathfrak{S}^j)$ feasible for $\mathcal{FA}(\gamma)$ naturally induces a randomized, stationary, history-independent policy, which we call the FA-$\boldsymbol{z}$ policy, denoted as $\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z})$. Note that, for any $\gamma>0$, $\mathcal{FA}(\gamma)$s are equivalent to each other. So the objective function value associated with $\boldsymbol{z}$, $\mathcal{V}_{\mathsf{FA}}(\boldsymbol{z})$, is independent of γ as well. We denote the solution to $\mathcal{FA}(\gamma)$ as $\boldsymbol{z}^*$, and the associated optimal objective function value as $\mathcal{V}_{\mathsf{FA}}^* = \mathcal{V}_{\mathsf{FA}}(\boldsymbol{z}^*)$.

Indeed, there are intrinsic connections between the optimal target problem $(\mathcal{OTP})$ and the FA convex program $\mathcal{FA}(\gamma)$. We can always construct a feasible (resp. optimal) click-through target vector in $(\mathcal{OTP})$ from any feasible (resp. optimal) probabilities in $\mathcal{FA}(\gamma)$. For any $\boldsymbol{z}$ feasible to $\mathcal{FA}(\gamma)$, we define $\hat{\boldsymbol{\alpha}}(\boldsymbol{z})\in\mathbb{R}^{(n+1)m}$, where $\hat{\alpha}_i^j(\boldsymbol{z}) = \sum_{i\in S\in\mathfrak{S}^j} p^j z^j(S)\phi_i^j(S)$.

LEMMA 1. *Assume that $\boldsymbol{z}$ is feasible in $\mathcal{FA}(\gamma)$. We have $\hat{\boldsymbol{\alpha}}(\boldsymbol{z})$ is first-stage feasible and can be achieved by policy $\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z})$, i.e., $\mathbb{E}_{\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}),\tilde{j}(t),\mathsf{D}_y}[\tilde{y}_i^j(t|\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}))] = \hat{\alpha}_i^j(\boldsymbol{z})$. Furthermore, $\mathcal{V}_{\mathsf{CT}}(\hat{\boldsymbol{\alpha}}(\boldsymbol{z})) = \mathcal{V}_{\mathsf{FA}}(\boldsymbol{z})$. In particular, $\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)$ is an optimal solution to $(\mathcal{OTP})$ with $\mathbb{E}_{\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*),\tilde{j}(t),\mathsf{D}_y}[\tilde{y}_i^j(t|\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*))] = \hat{\alpha}_i^j(\boldsymbol{z}^*)$.*

We are now ready to demonstrate that, as the problem size γ scales to infinity, the original problem $\mathcal{OP}(\gamma)$, the convex optimal target problem $(\mathcal{OTP})$, and the FA convex program $\mathcal{FA}(\gamma)$ all have the same optimal (expected) per-customer FAV, which is also identical to the one generated by the DWO-$\boldsymbol{\alpha}^*$ policy and the one generated by the FA-$\boldsymbol{z}^*$ policy in $\mathcal{OP}(\gamma)$. Formally, the following proposition establishes these equivalences and implies our main result.

PROPOSITION 2. *The following inequalities hold:*

$$\mathcal{V}^*_{\mathsf{CT}} \geq \mathcal{V}^*_{\mathsf{FA}} = \lim_{\gamma\uparrow+\infty} \mathcal{V}(\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*)|\gamma) = \lim_{\gamma\uparrow+\infty} \mathcal{V}^*(\gamma) \geq \lim_{\gamma\uparrow+\infty} \mathcal{V}(\tilde{\pi}_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma) = \mathcal{V}^*_{\mathsf{CT}} \tag{18}$$

Therefore, all inequalities in (18) *hold as equalities.*

Theorem 3 follows immediately from (18). The proof of Proposition 2 relies on a careful application of Theorem 2, which shows that the DWO-$\boldsymbol{\alpha}^*$ policy achieves the optimal FAV of the click-through target optimization (i.e., the last equality of (18)). Note that the first and second equalities in (18) generalize Proposition 1 of Liu and Van Ryzin (2008) to our setting with algorithmic fairness, personalized offer-sets, and click-through requirements, whereas the other (in)equalities thereof are new to the literature.

5.3. Discussions

Computational efficiency. Our proposed DWO-$\boldsymbol{\alpha}^*$ policy involves two steps. The first step solves the optimal target problem ($\mathcal{OTP}$) to obtain the optimal click-through target vector $\boldsymbol{\alpha}^*$ offline, and the second step implements the second-stage DWO display procedure online. We now discuss the computational efficiency of the two steps separately, starting from the second-stage online implementation.

Second-stage online implementation. To implement Algorithm 1 online, given any click-through target vector $\boldsymbol{\alpha}$, the bottleneck is to solve a single-period offer-set optimization problem (12) upon the arrival of each customer t. Standard results in the assortment-optimization literature suggest that if customers follow a wide range of commonly used choice models—such as the independent (Feldman et al. 2021), MNL (Rusmevichientong et al. 2010, Davis et al. 2013), generalized attraction (Gallego et al. 2015), and nested MNL (Davis et al. 2014) models—the personalized offer-set optimization (12) can be solved efficiently. Therefore, the second-stage online implementation of the DWO-$\boldsymbol{\alpha}$ is computationally efficient as long as the single-period offer-set optimization is tractable, which is generally the case for choice models commonly used in practice.

First-stage convex program. Section 4.2 and Appendix F show that if the customer choices follow the MNL, independent, and generalized attraction models, the optimal target problem ($\mathcal{OTP}$) can be greatly simplified to a convex program with a few linear constraints, which can be solved efficiently in general. We emphasize that the MNL, independent, and generalized attraction models are all widely used in practice. If customers follow a general choice model, Lemma 1 implies that ($\mathcal{OTP}$) shares the same computational complexity as the FA convex program $\mathcal{FA}(\gamma)$. To see this, note that for any solution to $\mathcal{FA}(\gamma)$, $\boldsymbol{z}^*$, we can construct a click-through target vector $\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)$ that is feasible and optimal for ($\mathcal{OTP}$). Therefore, as long as the auxiliary FA convex program $\mathcal{FA}(\gamma)$ is computationally tractable, we can efficiently obtain an optimal solution to $(\mathcal{OTP})$ as well. In general, the DWO-$\boldsymbol{\alpha}^*$ policy solves $(\mathcal{OTP})$ offline only once at the beginning of planning horizon, which is generally tractable in most applications. Indeed, Table 1 in Section 6 show that our DWO policy is much more scalable than the FA-based benchmarks commonly adopted in the literature.

If the auxiliary FA convex program $\mathcal{FA}(\gamma)$ is intractable, obtaining an *optimal* click-through target vector $\boldsymbol{\alpha}^*$ may be prohibitive. However, we can still identify a *feasible* click-through target vector $\boldsymbol{\alpha}$ by applying Theorem 1 and binary search ($\iota\boldsymbol{\alpha}$ is always feasible for a sufficiently small $\iota > 0$, so one can find a feasible $\iota\boldsymbol{\alpha}$ through a binary search on ι). As discussed above, once a feasible $\boldsymbol{\alpha}$ is found, the second-stage online implementation of the DWO-$\boldsymbol{\alpha}$ policy should be tractable, as long as the single-period offer-set optimization (12) is. Furthermore, the DWO-$\boldsymbol{\alpha}$ policy achieves the same asymptotic FAV as $\mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha})$, as shown in the following proposition.

PROPOSITION 3. *If $\boldsymbol{\alpha}$ is first-stage feasible, i.e., it is a feasible solution to $(\mathcal{OTP})$, then we have:*

$$\lim_{\gamma\uparrow+\infty} \mathcal{V}(\tilde{\pi}_{\mathsf{DWO}}(\boldsymbol{\alpha})|\gamma) = \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}). \tag{19}$$

The key implication from Proposition 3 is that, in the asymptotic regime, the second-stage online implementation of DWO display will not incur any additional optimality loss on top of that from a feasible suboptimal click-through target vector in the convex optimal target problem.

Comparison with existing algorithms. It is useful to compare the DWO policy with relevant algorithms in the existing literature.

Debt-weighted resource-allocation algorithms. As discussed above, the objective of existing debt-weighted algorithms (e.g, Zhong et al. 2017, Jiang et al. 2019, Lyu et al. 2019) is to allocate a centralized resource to satisfy some *feasible* and *exogenous* service-level constraints. The goal of the DWO policy, however, is to maximize the FAV of an online advertising system so that the click-through target vector, which is the counterpart of the service-level constraints in our setting, is *endogenized* in the first-stage of the algorithm. Because we have such a different objective for our policy, we develop a two-stage reformulation to implement the algorithm and take a different path for its analysis, which relies on establishing the (asymptotic) equivalence of different formulations of the problem. Another key difference between our DWO policy and other debt-weighted resource-allocation algorithms is that whereas those algorithms can freely control the allocation and consumption of the resources, our policy has to handle the additional complexity of customers' stochastic choice behaviors, which introduces another layer of challenge to controlling the debt process.

FA convex program heuristics. In our setting with a nonlinear fairness term, the standard LP-based heuristics with or without re-solving (e.g., Liu and Van Ryzin 2008, Jasin and Kumar 2012, Bumpensanti and Wang 2020) applied to linear rewards should be extended to similar heuristics based on FA convex programs (e.g., $\mathcal{FA}(\gamma)$). A core advantage of the DWO policy over the family of FA heuristics is that, by the nature of the algorithm to assign a higher weight to an ad-customer pair with a larger debt, the click-through and, thus, the reward process will follow a mean-reverting pattern. Therefore, compared with the FA or FA-resolving heuristics, our DWO algorithm can deplete the budget of each ad more smoothly throughout the horizon, which is highly desirable for the advertising business in practice. For example, Google recommends a "standard" ad delivery method for most advertisers, especially those with a low budget, to avoid exhausting their budgets early.[8] Under the "standard" delivery, each advertisement can reach customers evenly throughout the day. Furthermore, with the cardinality constraint on the displayed offer-sets, one difficulty using FA or its resolving variants is that the number of variables (i.e., the probability of each offer-set for all customer types) quickly explodes as the number of products increases, even when the choice model is restricted to MNL. The DWO policy has a better scalability than those FA-based heuristics. These advantages of our policy are also confirmed by our numerical comparisons in Section 6.

Inventory-balancing policy. Inventory balancing is another family of algorithms to address the personalized-assortment optimization problem with inventory constraints (e.g., Golrezaei et al. 2014). This policy uses the remaining inventory to reweight the value of each product. The inventory-balancing policy is difficult, if not impossible, to adapt to our setting because on one hand it is challenging for this policy to handle the click-through requirements $\{\eta_i^{\mathcal{C}} : i \in \bar{\mathcal{N}}, \mathcal{C} \in \mathfrak{K}_i\}$ and, on the other hand, it is hard to incorporate the nonlinear fairness metric into the inventory-balanced offer-set optimization problem upon the arrival of each customer. Our DWO policy circumvents these two challenges under our two-stage framework within which the personalized offer-set optimization problem is reduced to a standard single-period problem with a reward linear in the number of click-throughs (12). For completeness, we numerically compare our DWO policy with the inventory-balancing benchmark in the setting without click-through requirements (i.e., $\mathfrak{K}_i = \emptyset$ for all i) and fairness concerns (i.e., $\lambda = 0$) in Appendix I. The numerical results demonstrate that our policy outperforms the inventory-balancing benchmark for all the problem instances examined when the demand-to-supply ratio is not too large. When the demand-to-supply ratio is large, our policy performs fairly well, achieving an average of more than 99% of the theoretical upper-bound in all problem instances.

6. Numerical Experiments

In this section, we numerically evaluate our DWO policy for ad-allocation optimization, benchmarked against three FA-based heuristics. The first benchmark is the stationary history-independent policy induced by the optimal solution $\boldsymbol{z}^*$ to $\mathcal{FA}(\gamma)$ (see, also, Liu and Van Ryzin 2008), denoted as the

[8] See `https://support.google.com/google-ads/answer/2404248?hl=en` for more details.

fluid-approximation policy or the FA-$\boldsymbol{z}^*$ policy.[9] The second benchmark is a re-solving version of the FA policy, denoted as the *FA re-solving policy* or the FA-R policy, which re-solves the FA convex program at evenly spaced time epochs based on the remaining budgets and click-through requirements (see, e.g., Jasin and Kumar 2012). Finally, the third benchmark is an refined version of the FA-R policy, denoted as the *FA-infrequent re-solving policy* or the FA-I-R policy, under which the resolving time epochs are more carefully designed and are infrequent/sparse at the beginning of the ad campaign (see, e.g., Bumpensanti and Wang 2020). The implementation details of the FA-R and FA-I-R policies are provided in Appendix G.

The key takeaways from our numerical experiments are summarized as follows: (a) Thanks to the mean-reverting nature of our algorithm, the DWO policy achieves better-than-expected performance in FAV, and it delivers much more stable FAVs (i.e., with lower standard deviations) than the FA-based benchmarks. This is because the debt process of the DWO policy directly steers the click-throughs of each ad from each customer type toward the respective optimal target throughout the ad campaign. (b) The click-through requirements are substantially more likely to be satisfied by our DWO policy than the FA-based heuristics. This further justifies the efficacy of our algorithm in the presence of click-through requirements. (c) Our modeling framework together with the DWO policy could help achieve promising algorithmic fairness without compromising the advertising efficiency much. Therefore, our approach is well-suited for tackling the aforementioned challenge to address such algorithmic discrimination/bias in ad-delivery optimization.

We consider an ad-allocation problem with $T = 1,000$ customers of five types and 50 ads. The customer-type distribution $(p^1, p^2, ..., p^5)$, where $\mathbb{P}[\tilde{j}(t) = j] = p^j$ and $\sum_j p^j = 1$, is generated from a five-dimensional Dirichlet distribution (see Appendix J for details). We assume the per-click value depends only on ad type but not on customer type, and we sample the per-click value of each ad $\{r_1, r_2, ..., r_{50}\}$ independently from a uniform distribution on the interval $[10, 50]$. We model the click-through behavior of the customers using MNL, i.e., for $i \in S \in \mathfrak{S}^j$, $\phi_i^j(S)$ is given by (14). Each ad/customer-type pair is associated with an attraction index v_i^j. For ad i and customer-type j, let $v_i^j := \exp(u_i^j)$, where u_i^j is independently sampled from the uniform distribution on the interval $[0, 5]$. We set the cardinality constraint that the maximum size of an offer-set is 3, i.e., $|S(t)| \leq 3$ for each customer t. The fairness metric we use in the numerical studies is the GMD fairness (3) with $\lambda = 10$.

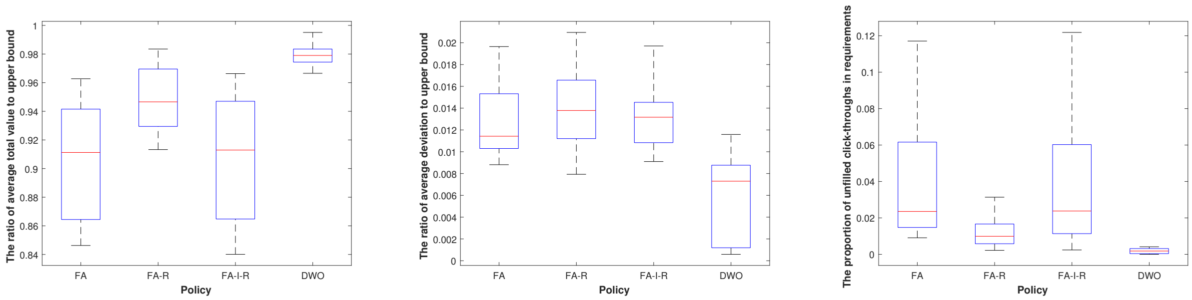
Our first set of numerical studies is based on problem instances generated by systematically varying two focal parameters: (a) the concentration parameter (CP) associated with the proportion of each customer type, and (b) the loading factor (LF), defined as the ratio between the total expected demand and total supply. Specifically, the concentration parameters are determined by the parameters

[9] When there is no confusion from the context, we drop $\boldsymbol{z}^*$ and abbreviate it as the FA policy.

of the Dirichlet distribution that we use to generate $(p^1, p^2, \cdots, p^5)$. The loading factor is the ratio between the total user traffic and the total affordable traffic with the budgets of the advertisers. (See Appendix J for detailed explanations of the CP parameter.) As is clear from their definitions, CP measures the uniformity of the customer-type distribution, while LF measures the tightness of the ad budget. The higher the CP, the more uniform the distribution of customer types; the higher the LF, the tighter the budget constraint for the ad campaigns. In our experiments, we vary CP in the set $\{0.1, 1, 10, 100\}$ and LF in the set $\{0.5, 0.75, 1, 1.25, 1.5\}$. For each problem instance, we solve the problem $(\mathcal{OTP}-\mathcal{MNL})$ with $\lambda = 0$ and without click-through requirements to obtain the solution-optimal targets $\boldsymbol{\alpha}^*$. A click-through requirement $\eta_i^{\{j\}}$ is generated by multiplying α_i^{j*} with a random number independently sampled from the $[0,1]$ uniform distribution for each ad i and each customer type j. We generate 30 sample paths for each problem instance to evaluate the following performance metrics of interest: (1) the ratio between the expected FAV and its theoretical upper bound characterized by the solution to the first-stage convex program, $\mathcal{V}^*_{\mathsf{CT}}$; (2) the ratio between the standard deviation of FAV and $\mathcal{V}^*_{\mathsf{CT}}$; and (3) the average proportion of unfilled click-through requirements. We use the relative ratios (instead of the absolute values) to make the comparisons clear.

We report the numerical findings as box plots with respect to all problem instances in Figure 2, which clearly illustrates the advantages of our algorithm over the benchmarks in various dimensions. Specifically, Figure 2(a) demonstrates that our DWO algorithm consistently outperforms all the FA-based benchmarks by delivering higher values in the total objective. Meanwhile, Figure 2(b) shows that the variability of FAV is much lower under our policy than those under the benchmarks. Finally, Figure 2(c) shows that the DWO algorithm significantly reduces the proportion of unfilled click-through requirements. In short, our proposed DWO algorithm not only generates higher FAV than the FA-based benchmarks do but also reduces the variability of FAV and increases the fill-rate of the click-through requirements.

To understand why our DWO algorithm enjoys the great performance illustrated in Figure 2, we also plot the 0.1-, 0.5- (i.e., median) and 0.9- quantiles of the click-through sample-paths of the highest-value ad for the four approaches we studied (under the problem instance $LF = 1$ and $CP = 100$) in Figure 3. As is clear from our numerical experiments, although all four policies deplete the ad's entire budget for more than 50% of sample paths, the variability of the click-through sample paths (equivalently, the budget-depleting process) through the entire time horizon under the FA, FA-R, and FA-I-R algorithms are much higher than our DWO policy. Furthermore, the FA-based approaches all run out of budget long before the end of the ad campaign, while our DWO policy exhausts the budget only toward the very end.



(a) Total Fairness-Adjusted Value

(b) Standard Deviation of FAV

(c) Unfilled Click-Throughs in Requirements

Figure 2 Comparison Between DWO-Based and FA-Based Benchmarks

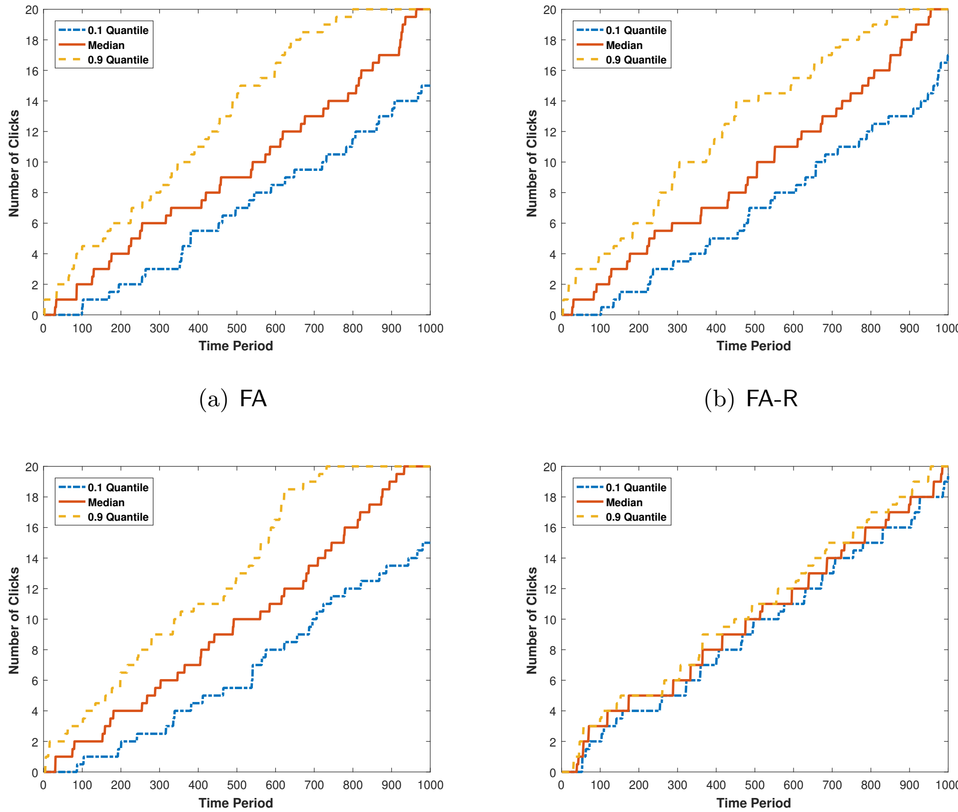
We highlight that such smooth budget depletion of our proposed algorithm should be credited to their mean-reverting pattern driven by the fact that the offer-set displayed in each period is prescribed in accordance with the “debts” owed by the algorithm to the optimal click-through targets.[10] Smooth budget depletion is a highly desirable property for advertisers of online advertising platforms in practice—Facebook has even built some API tools for its clients to pace their ad delivery and smooth their budget depletion.[11] Therefore, from the practicality perspective, our DWO algorithm may, appealingly, help advertisers and advertising platforms to achieve smoother budget depletion.

In addition to obtaining a better performance in most of the cases we examine and much smoother depletion of ad budgets, our DWO algorithm is more scalable and efficient in both time and space complexities. We carried out our numerical studies by varying the offer-set size constraint K from 2 to 5, with other model primitives identical to those of the experiments in Section 6. We conducted the experiment by using Gurobi 9.1 within MATLAB 9.0 on a 2.20GHz Intel Core i5-5200 CPU with 8 GB of RAM. Table 1 shows that the average computation time of finding optimal click-through goals (i.e., solving the convex program ($\mathcal{OTP}$)) is approximately 0.01s regardless of the value of K, but the FA policy is much more time-consuming (FA-R and FA-I-R are, of course, even slower). In addition, increasing K means exponentially more possible offer-sets for the FA, FA-R, and FA-I-R policies, so much more memory is needed in this case. Table 1 shows that the case of $K=5$ may even incur an “out of memory” error for the FA benchmark. In short, our algorithms enjoy higher scalability than the FA-based benchmarks.

Next, we demonstrate the efficiency-fairness trade-off by varying the parameter λ in our setting. Specifically, we consider the problem instance with $\lambda \in \{10^{-i_\lambda} : i_\lambda = -1 + 0.1 \times (i-1), i = 1, 2, ..., 31\} \cup$

[10] In Appendix H, we regress the click-through on the per-period debt for all four algorithms. A high per-period debt of an ad-customer pair has a much stronger impact on the potential click-throughs under our DWO policy compared to the FA, FA-R, and FA-I-R algorithms.

[11] See https://developers.facebook.com/docs/marketing-api/bidding/overview/pacing-and-scheduling.



(a) FA
(b) FA-R
(c) FA-I-R
(d) DWO

Figure 3 The 0.1 Quantiles, Medians, and 0.9 Quantiles in 30 Sample Paths Over Time of Click-Numbers of the Highest-Reward Advertisement With $LF = 1$, and $CP = 100$

Policy	$K = 2$	$K = 3$	$K = 4$	$K = 5$
DWO	0.003s	0.005s	0.005s	0.004s
FA	0.050s	0.421s	4.381s	out of memory

Table 1 The Comparison of Average Solving Time

$\{0\}$. We plot the relationship between the efficiency and Gini fairness in Figure 4 for different values of λ, where the x-axis (resp. y-axis) is the ratio between the expected efficiency (resp. expected GMD fairness) with respect to λ and that with respect to $\lambda = 0$ (i.e., the system is purely efficiency-driven). Our numerical results reveal the trade-off between efficiency and fairness. Importantly, we find that introducing the fairness term in the objective function could substantially reduce the algorithmic bias without much compromising the advertising efficiency. For example, a 1% (resp. 5%) optimality gap in efficiency could reduce about 50% (resp. 90%) of the algorithmic bias.

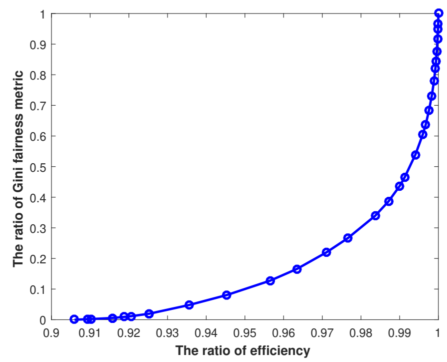


Figure 4 The Trade-Off Between the Optimal Efficiency and Gini Fairness

7. Conclusion

The allocation of customer traffic to different ads is a crucial operations decision for online e-commerce platforms to optimize their advertising business. The emerging advocacy for algorithmic fairness of online ad delivery has posed additional challenges for the design of ad-allocation policy. This paper proposes a general model and an associated efficient algorithm to study optimal ad allocation under customer choices and algorithmic fairness. Although the original online ad-allocation problem is intractable, we reformulate it into a two-stage stochastic program, and we demonstrate their asymptotic equivalence. Furthermore, we propose a simple and effective algorithm, referred to as the debt-weighted offer-set policy, which is provably optimal to achieve the maximum FAV from advertising in the asymptotic regime. Furthermore, the proposed algorithm gives rise to the mean-reverting pattern of the budget consumption process and, therefore, achieves smoother budget depletion, which is highly desirable from a practical perspective. Our algorithm also helps substantially improve the fairness of ad allocation for a platform without compromising its efficiency much.

References

Ali, Muhammad, Piotr Sapiezynski, Miranda Bogen, Aleksandra Korolova, Alan Mislove, Aaron Rieke. 2019. Discrimination through optimization: How facebook's ad delivery can lead to biased outcomes. *Proceedings of the ACM on Human-Computer Interaction* **3**(CSCW) 1–30.

Alptekinoğlu, Aydın, Arunava Banerjee, Anand Paul, Nikhil Jain. 2013. Inventory pooling to deliver differentiated service. *Manufacturing & Service Operations Management* **15**(1) 33–44.

Atkinson, Anthony B. 1975. The economics of inequality .

Balseiro, Santiago R, Jon Feldman, Vahab Mirrokni, Shan Muthukrishnan. 2014. Yield optimization of display advertising with ad exchange. *Management Science* **60**(12) 2886–2907.

Barocas, Solon, Moritz Hardt, Arvind Narayanan. 2017. Fairness in machine learning. *Nips tutorial* **1** 2017.

Bernstein, Fernando, A Gürhan Kök, Lei Xie. 2015. Dynamic assortment customization with limited inventories. *Manufacturing & Service Operations Management* **17**(4) 538–553.

Bernstein, Fernando, Sajad Modaresi, Denis Sauré. 2019. A dynamic clustering approach to data-driven assortment personalization. *Management Science* **65**(5) 2095–2115.

Bertsimas, Dimitris, Vivek F Farias, Nikolaos Trichakis. 2012. On the efficiency-fairness trade-off. *Management Science* **58**(12) 2234–2250.

Bumpensanti, Pornpawee, He Wang. 2020. A re-solving heuristic with uniformly bounded loss for network revenue management. *Management Science* **66**(7) 2993–3009.

Caro, Felipe, Victor Martínez-de Albéniz, Paat Rusmevichientong. 2014. The assortment packing problem: Multiperiod assortment planning for short-lived products. *Management Science* **60**(11) 2701–2721.

Celis, Elisa, Anay Mehrotra, Nisheeth Vishnoi. 2019. Toward controlling discrimination in online ad auctions. *International Conference on Machine Learning*. PMLR, 4456–4465.

Chen, Xi, Will Ma, David Simchi-Levi, Linwei Xin. 2016. Assortment planning for recommendations at checkout under inventory constraints. *Available at SSRN 2853093* .

Cheung, Wang Chi, David Simchi-Levi. 2017. Thompson sampling for online personalized assortment optimization problems with multinomial logit choice models. *Available at SSRN 3075658* .

Choi, Hana, Carl F Mela, Santiago R Balseiro, Adam Leary. 2020. Online display advertising markets: A literature review and future directions. *Information Systems Research* .

Davis, James, Guillermo Gallego, Huseyin Topaloglu. 2013. Assortment planning under the multinomial logit model with totally unimodular constraint structures. *Work in Progress* .

Davis, James M, Guillermo Gallego, Huseyin Topaloglu. 2014. Assortment optimization under variants of the nested logit model. *Operations Research* **62**(2) 250–273.

Dong, Lingxiu, Duo Shi, Fuqiang Zhang. 2017. 3D printing vs. traditional flexible technology: Implications for operations strategy. *Available at SSRN 2847731* .

Eppen, Gary D. 1979. Note—effects of centralization on expected costs in a multi-location newsboy problem. *Management science* **25**(5) 498–501.

Feldman, Jacob, Dennis J Zhang, Xiaofei Liu, Nannan Zhang. 2021. Customer choice models vs. machine learning: Finding optimal product displays on alibaba. *Operations Research* .

Gallego, Guillermo, Anran Li, Van-Anh Truong, Xinshang Wang. 2016. Online personalized resource allocation with customer choice. Tech. rep., Working Paper. http://arxiv. org/abs/1511.01837 v1.

Gallego, Guillermo, Richard Ratliff, Sergey Shebalov. 2015. A general attraction model and sales-based linear program for network revenue management under customer choice. *Operations Research* **63**(1) 212–232.

Golrezaei, Negin, Hamid Nazerzadeh, Paat Rusmevichientong. 2014. Real-time optimization of personalized assortments. *Management Science* **60**(6) 1532–1551.

Hao, Xiaotian, Zhaoqing Peng, Yi Ma, Guan Wang, Junqi Jin, Jianye Hao, Shan Chen, Rongquan Bai, Mingzhou Xie, Miao Xu, Zhenzhe Zheng, Chuan Yu, HAN LI, Jian Xu, Kun Gai. 2020. Dynamic knapsack optimization towards efficient multi-channel sequential advertising. *ICML 2020* .

Hojjat, Ali, John Turner, Suleyman Cetintas, Jian Yang. 2017. A unified framework for the scheduling of guaranteed targeted display advertising under reach and frequency requirements. *Operations Research* **65**(2) 289–313.

Imana, Basileal, Aleksandra Korolova, John Heidemann. 2021. Auditing for discrimination in algorithms delivering job ads. *Proceedings of the Web Conference 2021*. 3767–3778.

Jasin, Stefanus, Sunil Kumar. 2012. A re-solving heuristic with bounded revenue loss for network revenue management with customer choice. *Mathematics of Operations Research* **37**(2) 313–345.

Jiang, Jiashuo, Shixin Wang, Jiawei Zhang. 2019. Achieving high individual service-levels without safety stock? optimal rationing policy of pooled resources. Working Paper.

Kallus, Nathan, Xiaojie Mao, Angela Zhou. 2021. Assessing algorithmic fairness with unobserved protected class using data combination. *Management Science* .

Kallus, Nathan, Madeleine Udell. 2020. Dynamic assortment personalization in high dimensions. *Operations Research* .

Kumar, Amit, Jon Kleinberg. 2000. Fairness measures for resource allocation. *Proceedings 41st annual symposium on foundations of computer science*. IEEE, 75–85.

Lambrecht, Anja, Catherine Tucker. 2019. Algorithmic bias? an empirical study of apparent gender-based discrimination in the display of stem career ads. *Management science* **65**(7) 2966–2981.

Lejeune, Miguel A, John Turner. 2019. Planning online advertising using gini indices. *Operations Research* **67**(5) 1222–1245.

Liu, Qian, Garrett Van Ryzin. 2008. On the choice-based linear programming model for network revenue management. *Manufacturing & Service Operations Management* **10**(2) 288–310.

Luce, R Duncan. 2012. *Individual choice behavior: A theoretical analysis*. Courier Corporation.

Lyu, Guodong, Wang-Chi Cheung, Mabel C Chou, Chung-Piaw Teo, Zhichao Zheng, Yuanguang Zhong. 2019. Capacity allocation in flexible production networks: Theory and applications. *Management Science* **65**(11) 5091–5109.

Lyu, Guodong, Mabel C Chou, Chung-Piaw Teo, Zhichao Zheng, Yuanguang Zhong. 2021. Stochastic knapsack revisited: The service level perspective. *Operations Research* .

Ma, Will, Pan Xu. 2020. Group-level fairness maximization in online bipartite matching. *arXiv preprint arXiv:2011.13908* .

Microsoft. 2020. *Partner Incentives Co-op Guidebook: Business Policies for FY21*. Microsoft Corporation.

Nakamura, Atsuyoshi, Naoki Abe. 2005. Improvements to the linear programming based scheduling of web advertisements. *Electronic Commerce Research* **5**(1) 75–98.

Rusmevichientong, Paat, Zuo-Jun Max Shen, David B Shmoys. 2010. Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. *Operations Research* **58**(6) 1666–1680.

Shen, Huaxiao, Yanzhi Li, Youhua Chen, Kai Pan. 2021a. Integrated ad delivery planning for targeted display advertising. *Operations Research* .

Shen, Huaxiao, Yanzhi Li, Jingjing Guan, Geoffrey KF Tso. 2021b. A planning approach to revenue management for non-guaranteed targeted display advertising. *Production and Operations Management* **30**(6) 1583–1602.

Speicher, Till, Muhammad Ali, Giridhari Venkatadri, Filipe Nunes Ribeiro, George Arvanitakis, Fabrício Benevenuto, Krishna P Gummadi, Patrick Loiseau, Alan Mislove. 2018. Potential for discrimination in online targeted advertising. *Conference on Fairness, Accountability and Transparency*. PMLR, 5–19.

Swaminathan, Jayashankar M, Ramesh Srinivasan. 1999. Managing individual customer service constraints under stochastic demand. *Operations Research Letters* **24**(3) 115–125.

Turner, John. 2012. The planning of guaranteed targeted display advertising. *Operations research* **60**(1) 18–33.

Wang, Ruxian. 2012. Capacitated assortment and price optimization under the multinomial logit model. *Operations Research Letters* **40**(6) 492–497.

Yang, Jian, Erik Vee, Sergei Vassilvitskii, John Tomlin, Jayavel Shanmugasundaram, Tasos Anastasakos, Oliver Kennedy. 2012. Inventory allocation for online graphical display advertising using multi-objective optimization. *ICORES* **12** 293–304.

Ye, Zikun, Dennis Zhang, Heng Zhang, Renyu Philip Zhang, Xin Chen, Zhiwei Xu. 2020. Cold start on online advertising platforms: Data-driven algorithms and field experiments. *Available at SSRN* .

Young, H Peyton, R Mark Isaac. 1995. Equity: In theory and practice. *Journal of Economic Literature* **33**(1) 210–210.

Zhong, Yuanguang, Zhichao Zheng, Mabel C Chou, Chung-Piaw Teo. 2017. Resource pooling and allocation policies to deliver differentiated service. *Management Science* **64**(4) 1555–1573.

Online Appendices

A. Summary of Notations

Notation	Description
T	Total number of customers
$\mathcal{N}$	Set of ad campaigns
n	Number of ad campaigns
$\bar{\mathcal{N}}$	Set of all ad campaigns and the organic recommendation
B_i	Total budget of ad campaign i
b_i	Bid price of ad campaign i per click-through
$j(t)$ ($\tilde{j}(t)$)	Realized (random) type of customer t
$\mathcal{M}$	Set of customer types
m	Number of customer types
p^j	Probability of a customer being type j
$S(t)$ ($S^j(t)$)	Offer-set displayed to customer t (if the customer type is j)
$\mathfrak{S}^j$	Collection of all possible offer-sets for type-j customers including ad targeting info
$y_i^j(t)$ ($\tilde{y}_i^j(t)$)	Realized (random) number of click-throughs by a type-j customer on ad i in type t
$\bar{y}_i^j$	Per-period value of $y_i^j(t)$
$\phi_i^j(S)$	Expected value of $\tilde{y}_i^j(t)$ conditioned on $S(t)=S$
D_y	Click-through distribution
$\mathcal{L}_i$	Set of customer types targeted by ad campaign i
$\eta_i^{\mathcal{C}}$	Required click-throughs for customer-type set $\mathcal{C}$ on ad campaign i
r_i^j	Value of each click of ad campaign i by a type-j customer
$F(\boldsymbol{y})$	Fairness metric
$\mathcal{H}_{t-1}$	Realized history until the start of time t
Π (π)	Set of (one) deterministic policies (policy)
$\tilde{\Pi}$ ($\tilde{\pi}$)	Set of (one) randomized policies (policy)
Π_{s}	Set of deterministic stationary history-independent policies
$\tilde{\Pi}_{\mathsf{s}}$	Set of randomized stationary history-independent policies
($\mathcal{OP}$)	Original stochastic program
$\mathcal{V}^*$	Optimal FAV of the original stochastic program
α_j^i	Target for the per-period number of click-throughs of ad i from type-j customers
$\mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha})$	FAV of the click-through target vector $\boldsymbol{\alpha}$
($2\mathcal{SSP}$)	Two-stage stochastic program
θ_i^j	Dual variable associated with satisfying the click-through target ad i from type-j customers
($\mathcal{OTP}$)	Reformulated optimal target problem
$\boldsymbol{\alpha}^*$	Solution to ($\mathcal{OTP}$)
K	Maximum size of an offer-set
DWO-$\boldsymbol{\alpha}$	DWO policy with click-through target vector $\boldsymbol{\alpha}$
$d_i^j(t)$	Debt of the click-throughs from type-j customers on ad i in time t
γ	Scale of a problem
$\mathcal{Q}(\gamma)$	Family of ad-allocation problems given scaling parameter γ
$\mathcal{V}(\tilde{\pi}\|\gamma)$	Expected value of policy $\tilde{\pi}$ for $\mathcal{Q}(\gamma)$
$\mathcal{FA}(\gamma)$	Fluid-approximation convex program with scale γ
$\mathcal{V}_{\mathsf{FA}}(\boldsymbol{z})$	Objective value function of $\mathcal{FA}(\gamma)$

Table 2 Summary of Notations

B. Dynamic Program Formulation

In this section, we formulate the original stochastic program $(\mathcal{OP})$ as a DP. Specifically, we define $Y_i^j(t) := \sum_{\tau=1}^{t-1} y_i^j(\tau)$ as the accumulative number of click-throughs until the start of time t, and

$$\begin{aligned}
\mathbb{V}_t(\boldsymbol{Y}(t)) := \max_{\tilde{\pi}\in\tilde{\Pi}} \mathbb{E}_{\tilde{\pi},\tilde{\boldsymbol{j}},\mathsf{D}_y} &\left[\sum_{\tau=t}^{T} \sum_{i\in\bar{\mathcal{N}}} \sum_{j\in\mathcal{L}_i} r_i^j \tilde{y}_i^j(t|\tilde{\pi}_t, \mathcal{H}_{t-1}) + T\lambda F(\bar{\mathbf{y}}(\tilde{\pi})) \middle| \boldsymbol{Y}(t) \right] \\
\text{s.t. } &\sum_{\tau=t}^{T} \sum_{j\in\mathcal{M}} b_i \tilde{y}_i^j(t|\tilde{\pi}_t, \mathcal{H}_{t-1}) \le B_i - \sum_{j\in\mathcal{M}} b_i Y_i^j(t), \text{ } \textit{almost surely} \text{ for each } i\in\mathcal{N}, \\
&\sum_{\tau=t}^{T} \sum_{j\in\mathcal{C}} \tilde{y}_i^j(t|\tilde{\pi}_t, \mathcal{H}_{t-1}) \ge \eta_i^{\mathcal{C}} - \sum_{j\in\mathcal{C}} Y_i^j(t), \text{ } \textit{almost surely} \text{ for each } i\in\bar{\mathcal{N}} \text{ and } \mathcal{C}\in\mathfrak{K}_i.
\end{aligned} \tag{20}$$

Hence, $\mathbb{V}_t(\boldsymbol{Y}(t))$ is the maximum expected FAV given that the number of accumulative click-throughs at the beginning of time t is $\boldsymbol{Y}(t) := (Y_i^j(t) : i\in\bar{\mathcal{N}}, j\in\mathcal{M})$.

To formulate the DP, we first specify the boundary/terminal value function $\mathbb{V}_{T+1}(\boldsymbol{Y}(T+1))$. To this end, we define

$$\mathcal{Y} := \left\{ \boldsymbol{Y}(T+1) \in \mathbb{R}_+^{(n+1)m} : b_i \sum_{j\in\mathcal{M}} Y_i^j(T+1) \le B_i \text{ for all } i\in\bar{\mathcal{N}}, \text{ and } \sum_{j\in\mathcal{C}} Y_i^j(T+1) \ge \eta_i^{\mathcal{C}}, \text{ for all } i\in\bar{\mathcal{N}} \text{ and } \mathcal{C}\in\mathfrak{K}_i \right\}$$

as the feasible region for the accumulative number of click-throughs for the entire planning horizon, $\boldsymbol{Y}(T+1)$. The boundary value function is defined as follows:

$$\mathbb{V}_{T+1}(\boldsymbol{Y}(T+1)) = \begin{cases} T\lambda F\left(\frac{\boldsymbol{Y}(T+1)}{T}\right), & \text{if } \boldsymbol{Y}(T+1)\in\mathcal{Y}, \\ -\bar{M}, & \text{otherwise}, \end{cases} \tag{21}$$

where $\bar{M}$ is a sufficiently large positive number that is far bigger than $\mathcal{V}^*$ (e.g., $\bar{M} := C\cdot\max\{\mathcal{V}^*, 1\}$, where $C>0$ is a very large positive number).

By the standard backward induction argument, we are now ready to write the Bellman equation to evaluate $\mathbb{V}_t(\boldsymbol{Y}(t))$ in (20):

$$\mathbb{V}_t(\boldsymbol{Y}(t)) = \sum_{j\in\mathcal{M}} p_j \max_{S(t)\in\mathfrak{S}^j} \mathbb{E}_{\tilde{\boldsymbol{y}}(t)} \left[\sum_{i\in\bar{\mathcal{N}}} r_i^j \tilde{y}_i^j + \mathbb{V}_t(\boldsymbol{Y}(t) + \tilde{\boldsymbol{y}}(t)) \middle| S(t), \tilde{j}(t) = j \right]. \tag{22}$$

Therefore, the optimal FAV for the original problem $(\mathcal{OP})$ is

$$\mathcal{V}^* = \frac{\mathbb{V}_1(\mathbf{0})}{T}, \text{ where } \mathbf{0} := (0,0,...,0)' \in \mathbb{R}^{(n+1)m}.$$

Due to the curse of dimensionality, the above DP formulation of $(\mathcal{OP})$ is intractable even when m and n are just moderately large. Therefore, we resort to the two-stage formulation and the DWO algorithm to solve the ad-allocation optimization problem.

C. Derivation of the Gini Mean Difference Fairness

In this section, we derive the Gini Mean Difference metric (see, also, Lejeune and Turner 2019), which we also adopt in the numerical experiments.

Following Lejeune and Turner (2019), given the average click-through vector of ad j $\bar{\boldsymbol{y}}_i = (\bar{y}_i^1, y_i^2, ..., y_i^m)'$, we first define the GMD fairness for each ad i:

$$GMD_i(\bar{\boldsymbol{y}}_i) = \frac{2}{(\sum_{j\in\mathcal{L}_i} p^j)^2} \sum_{j,j'\in\mathcal{L}_i} p^j p^{j'} \left| \frac{\bar{y}_i^j}{p^j} - \frac{\bar{y}_i^{j'}}{p^{j'}} \right|, \tag{23}$$

where p^j is the proportion of type-j customers, and $\frac{\bar{y}_i^j}{p^j} = \frac{\sum_{t=1}^T \tilde{y}_i^j(t)}{p^j T}$ is the per-type-j customer click-throughs of ad i. Hence, the Gini coefficient of ad i is defined as follows:

$$G_i(\bar{\boldsymbol{y}}_i) = \frac{\left(\sum_{j\in\mathcal{L}_i} p^j\right) GMD_i(\bar{\boldsymbol{y}}_i)}{2\sum_{j\in\mathcal{L}_i} \bar{y}_i^j}.$$

We are now ready to define the GMD fairness as a weighted sum of the negative Gini coefficient of each ad:

$$F(\bar{\boldsymbol{y}}) = -\sum_{i\in\bar{\mathcal{N}}} k_i G_i(\bar{\boldsymbol{y}}_i) = -\sum_{i\in\bar{\mathcal{N}}} \frac{k_i}{\left(\sum_{j\in\mathcal{L}_i} \bar{y}_i^{j'}\right)\left(\sum_{j\in\mathcal{L}_i} p^j\right)} \sum_{j,j'\in\mathcal{L}_i} |p^{j'}\bar{y}_i^j - p^j \bar{y}_i^{j'}|,$$

where $k_i \geq 0$ is the weight of ad i according to its importance in the GMD fairness metric. Following Lejeune and Turner (2019), we choose $k_i = \sum_{j\in\mathcal{L}_i} \bar{y}_i^j$ which gives rise to our GMD fairness metric as Eq. (3).

D. Proof of Statements

We provide the proof of all the technical results in this section.

Proof of Theorem 1

The proof follows from the discussions before Eq. (11). □

Proof of Proposition 1

Before proving Proposition 1, we first state and prove a few auxiliary results. It is sometimes more convenient to use a binary variable representation of a deterministic stationary history-independent offer-set policy $\pi \in \Pi_{\mathsf{s}}$. More specifically, $\pi \in \Pi_{\mathsf{s}}$ can be equivalently represented by an $(n+1)m$-dimensional binary vector $\boldsymbol{x} = (x_i^j \in \{0,1\} : i \in \bar{\mathcal{N}}, j \in \mathcal{M})$, where $x_i^j = 1$ means that $i \in \pi(j)$, i.e., ad i is included in the offer-set displayed to a type j customer. Hence, $x_0^j = 1$ for all $j \in \mathcal{M}$. With a slight abuse of notation, we denote $\phi_i^j(\boldsymbol{x})$ as the expected click-throughs of a type j customer for ad i if the offer-set displayed to this customer is $S^j = \{i \in \bar{\mathcal{N}} : x_i^j = 1\}$. Under the MNL model, we have

$$\phi_i^j(\boldsymbol{x}) = \frac{v_i^j x_i^j}{1 + \sum_{i'\in\bar{\mathcal{N}}} v_{i'}^j x_{i'}^j}, \tag{24}$$

where $v_i^j > 0$ is the attractiveness of ad i to type-j customers (see, also, Eq. (14)). Denote the set of all plausible offer-set representation vectors as $\mathcal{X} \subset \{0,1\}^{(n+1)m}$, and the set of plausible offer-set representation vectors displayed to a type-j customer as $\mathcal{X}^j \subset \{0,1\}^{(n+1)}$. Applying Theorems 1 to the MNL choice model (24), we have the following corollary.

Corollary 1. *If customers follow the MNL click-through model (24), a click-through target vector $\boldsymbol{\alpha}$ is single-period feasible if and only if*

$$\max_{\boldsymbol{x}\in\mathcal{X}} \sum_{i\in\bar{\mathcal{N}}, j\in\mathcal{M}} \frac{p^j v_i^j \theta_i^j x_i^j}{1 + \sum_{i'\in\bar{\mathcal{N}}} v_{i'}^j x_{i'}^j} \geq \sum_{i\in\bar{\mathcal{N}}, j\in\mathcal{M}} \alpha_i^j \theta_i^j \text{ for all } \theta_i^j \geq 0 \ (i \in \bar{\mathcal{N}}, j \in \mathcal{M}). \tag{25}$$

Furthermore, (25) is equivalent to, for each $j \in \mathcal{M}$

$$\max_{\boldsymbol{x}^j\in\mathcal{X}^j} \sum_{i\in\bar{\mathcal{N}}} \frac{p^j v_i^j \theta_i^j x_i^j}{1 + \sum_{i'\in\bar{\mathcal{N}}} v_{i'}^j x_{i'}^j} \geq \sum_{i\in\bar{\mathcal{N}}} \alpha_i^j \theta_i^j \text{ for any } \theta_i^j \geq 0 \ (i \in \bar{\mathcal{N}}). \tag{26}$$

Proof of Corollary 1

Directly applying Theorem 1 to the MNL choice model implies that $\boldsymbol{\alpha}$ is feasible if and only if inequality (25) holds.

We now show that (25) implies (26). If (25) holds for any $\boldsymbol{\theta} \geq \mathbf{0}$, then it also holds for any $\boldsymbol{\theta} \geq \mathbf{0}$ with $\theta_i^{j'} = 0$ (for $j' \neq j$ and all $i \in \bar{\mathcal{N}}$). Therefore, (26) holds.

Finally, we show that if (26) holds for all j, (25) holds as well. Note that the left-hand side of (25) can be decomposed into independent parts as follows:

$$\max_{\boldsymbol{x}\in\mathcal{X}} \sum_{i\in\bar{\mathcal{N}},j\in\mathcal{M}} \frac{p^j v_i^j \theta_i^j x_i^j}{1+\sum\limits_{i'\in\bar{\mathcal{N}}} v_{i'}^j x_{i'}^j} = \max_{\boldsymbol{x}\in\mathcal{X}} \sum_{j\in\mathcal{M}} \sum_{i\in\bar{\mathcal{N}}} \frac{p^j v_i^j \theta_i^j x_i^j}{1+\sum\limits_{i'\in\bar{\mathcal{N}}} v_{i'}^j x_{i'}^j} = \sum_{j\in\mathcal{M}} \max_{\boldsymbol{x}^j\in\mathcal{X}^j} \sum_{i\in\bar{\mathcal{N}}} \frac{p^j v_i^j \theta_i^j x_i^j}{1+\sum\limits_{i'\in\bar{\mathcal{N}}} v_{i'}^j x_{i'}^j} \geq \sum_{j\in\mathcal{M}} \sum_{i\in\bar{\mathcal{N}}} \alpha_i^j \theta_i^j = \sum_{i\in\bar{\mathcal{N}},j\in\mathcal{M}} \alpha_i^j \theta_i^j,$$

for any $\boldsymbol{\theta} \geq \mathbf{0}$, where the inequality follows from (26). Therefore, that (25) holds is equivalent to that (26) holds for all j. This completes the proof of Corollary 1. □

Leveraging the structural properties of the MNL model, we can give a sharper and simpler characterization for the feasibility condition (as the solution to a linear program). The following lemma characterizes the feasibility condition for a click-through target vector $\boldsymbol{\alpha}$, taking into account the cardinality constraint that the size of an offer-set displayed to any customer is upper bounded by K, i.e., $|\tilde{S}(t)| \leq K$ for any customer t.

Lemma 2. *If customers follow the MNL click-through model (24) and the size of an offer-set cannot exceed K, we have $\boldsymbol{\alpha}$ is single-period feasible if and only if there exist $\boldsymbol{w} := (w_i^j : i \in \bar{\mathcal{N}}, j \in \mathcal{M})$ and $\boldsymbol{z} := (z^j : j \in \mathcal{M})$ that satisfy the following linear constraints*

$$\begin{aligned} & p^j v_i^j w_i^j \geq \alpha_i^j,\ w_i^j \leq z^j,\ 0 \leq w_i^j \leq 1 \quad \forall i,j \\ & \sum_{i=0}^{n} v_i^j w_i^j + z^j = 1,\ \sum_{i=0}^{n} w_i^j \leq K z^j,\ w_0^j = z^j \quad \forall j, \end{aligned} \tag{27}$$

where $z^j := 1/(1+\sum_{i'} v_{i'}^j x_{i'}^j)$ and $w_i^j := x_i^j z^j = x_i^j/(1+\sum_{i'} v_{i'}^j x_{i'}^j)$.

Proof of Lemma 2 A standard result in the assortment optimization literature postulates that the left-hand side of (26) is quasi-convex in $\boldsymbol{x}^j$ for all j, so there always exists a maximizer on the boundary of the feasible region. Thus, we can relax the binary constraint $x_i^j \in \{0,1\}$ to $x_i^j \in [0,1]$ in (26), which is therefore equivalent to

$$\max_{\boldsymbol{x}^j\in[0,1]^{(n+1)},\sum_i x_i^j \leq K} \sum_{i\in\bar{\mathcal{N}}} \frac{p^j v_i^j \theta_i^j x_i^j}{1+\sum_{i'\in\bar{\mathcal{N}}} v_{i'}^j x_{i'}^j} \geq \sum_{i\in\bar{\mathcal{N}}} \alpha_i^j \theta_i^j \text{ for all } \boldsymbol{\theta}^j \geq \mathbf{0} \text{ and } j \in \mathcal{M}. \tag{28}$$

We change the decision variable and define, for all $j \in \mathcal{M}$,

$$z^j := \frac{1}{1+\sum\limits_{i'} v_{i'}^j x_{i'}^j} \text{ and } w_i^j := x_i^j z^j = \frac{x_i^j}{1+\sum\limits_{i'} v_{i'}^j x_{i'}^j}.$$

Since $x_0^j = 1$ for all j, we have $w_0^j = z^j$ for all $j \in \mathcal{M}$. Then, we can rewrite (28) as, for any j,

$$\begin{aligned} \min_{\boldsymbol{\theta}\geq\mathbf{0}} & \left(\max_{w_i^j, z^j} \sum_{i\in\bar{\mathcal{N}}} p^j v_i^j w_i^j \theta_i^j - \sum_{i\in\bar{\mathcal{N}}} \alpha_i^j \theta_i^j \right) \geq 0 \\ \text{s.t. } & \sum_{i=0}^{n} v_i^j w_i^j + z^j = 1, \\ & \sum_{i\in\bar{\mathcal{N}}} w_i^j \leq K z^j, \\ & w_i^j \leq z^j \text{ for all } i \in \bar{\mathcal{N}}, \\ & w_0^j = z^j. \end{aligned} \tag{29}$$

By Sion's minimax theorem, we can exchange the maximization and minimization operators so that (29) is equivalent to, for any $j \in \mathcal{M}$:

$$\begin{aligned} \max_{\boldsymbol{w}^j, z^j} \min_{\boldsymbol{\theta}^j \geq 0} \ & \sum_{i=0}^{n} \theta_i^j (p^j v_i^j w_i^j - \alpha_i^j) \geq 0, \\ \text{s.t.} \ & \sum_{i=0}^{n} v_i^j w_i^j + z^j = 1, \\ & \sum_{i=0}^{n} w_i^j \leq K z^j, \\ & 0 \leq w_i^j \leq z^j, \ \text{for all } i \in \bar{\mathcal{N}}, \\ & w_0^j = z^j. \end{aligned} \tag{30}$$

Therefore, (30) holds if and only if there exist $\boldsymbol{w}^j$ and z^j such that all the constraints in (30) hold and $\sum_{i=0}^{n} \theta_i^j (p^j v_i^j w_i^j - \alpha_i^j) \geq 0$ holds for all $\boldsymbol{\theta}^j \geq \mathbf{0}$, which is equivalent to $p^j v_i^j w_i^j - \alpha_i^j \geq 0$ for all $i \in \bar{\mathcal{N}}$. Therefore, (30) is equivalent to that, for any $j \in \mathcal{M}$,

$$\begin{aligned} & p^j v_i^j w_i^j - \alpha_i^j \geq 0, \\ & \sum_{i=0}^{n} v_i^j w_i^j + z^j = 1, \\ & \sum_{i=0}^{n} w_i^j \leq K z^j, \\ & 0 \leq w_i^j \leq z^j, \ \text{for all } i \in \bar{\mathcal{N}}, \\ & w_0^j = z^j. \end{aligned} \tag{31}$$

That (31) holds for all $j \in \mathcal{M}$ is equivalent to that (27) holds. This completes the proof of Lemma 2. □

We now prove Proposition 1 itself. It suffices to show that, taking into account the cardinality constraint $|\tilde{S}(t)| \leq K$, the (*first-stage*) feasible region for the first-stage click-through target vector $\boldsymbol{\alpha}$ is given by the following linear constraints:

$$\mathcal{A}_{MNL} := \left\{ \boldsymbol{\alpha} \in \mathbb{R}_+^{(n+1)m} : \sum_{i'=1}^{n} \alpha_{i'}^j + \frac{(1+v_0^j)\alpha_i^j}{v_i^j} \leq p^j, \forall i, j, \ \text{and} \ \sum_{i=1}^{n} \alpha_i^j + \frac{1+v_0^j}{K-1} \sum_{i=1}^{n} \frac{\alpha_i^j}{v_i^j} \leq p^j, \forall j \right\}. \tag{32}$$

We first show that if (27) holds, then $\boldsymbol{\alpha} \in \mathcal{A}_{MNL}$. By the first inequality of (27), we have $v_i^j w_i^j \geq \frac{\alpha_i^j}{p^j}$ for all i and j. Plugging this into the first and second equalities of (27), we have

$$1 - z^j - z^j v_0^j = \sum_{i=1}^{n} v_i^j w_i^j \geq \sum_{i=1}^{n} \frac{\alpha_i^j}{p^j}.$$

Thus, by the first and second inequalities of (27), we have

$$\sum_{i'=1}^{n} \frac{\alpha_{i'}^j}{p^j} \leq 1 - z^j - z^j v_0^j \leq 1 - (1+v_0^j) w_i^j \leq 1 - \frac{(1+v_0^j)\alpha_i^j}{p^j v_i^j} \ \text{for any } i, j.$$

Rearranging the terms, we have

$$p^j \geq \sum_{i'=1}^{n} \alpha_{i'}^j + \frac{(1+v_0^j)\alpha_i^j}{v_i^j} \ \text{for any } i, j.$$

The first, second, and fourth inequalities of (27) imply that

$$\sum_{i=1}^{n}\frac{\alpha_i^j}{p^j v_i^j}\le\sum_{i=1}^{n}w_i^j\le(K-1)z^j=(K-1)\frac{1-\sum_{i=1}^{n}v_i^j w_i^j}{1+v_0^j}\le\frac{K-1}{1+v_0^j}\left(1-\sum_{i=1}^{n}\frac{\alpha_i^j}{p^j}\right)\text{ for all } j.$$

Rearranging the terms, we have

$$p^j\ge\sum_{i=1}^{n}\alpha_i^j+\frac{1+v_0^j}{K-1}\sum_{i=1}^{n}\frac{\alpha_i^j}{v_i^j}.$$

Therefore, if (27) holds, we have $\boldsymbol{\alpha}\in\mathcal{A}_{MNL}$.

Next, we show that if $\boldsymbol{\alpha}\in\mathcal{A}_{MNL}$, (27) holds. Given $\boldsymbol{\alpha}\in\mathcal{A}_{MNL}$, define

$$w_i^j=\frac{\alpha_i^j}{p^j v_i^j}\text{ for each } i\in\mathcal{N}\text{ and } j\in\mathcal{M},\text{ and } w_0^j=z^j=\frac{1}{1+v_0^j}\left(1-\sum_{i=1}^{n}\frac{\alpha_i^j}{p^j}\right).$$

To show (27), it suffices to show the first, second and fourth inequalities hold because the rest of the constraints hold trivially.

Since $p_j\ge\sum_{i=1}^{n}\alpha_i^j+\frac{1+v_0^j}{K-1}\sum_{i=1}^{n}\frac{\alpha_i^j}{v_i^j}$ for all j, we have

$$\sum_{i=1}^{n}w_i^j=\sum_{i=1}^{n}\frac{\alpha_i^j}{p^j v_i^j}=\frac{1}{p^j}\sum_{i=1}^{n}\frac{\alpha_i^j}{v_i^j}\le\frac{K-1}{1+v_0^j}\left(1-\sum_{i=1}^{n}\frac{\alpha_i^j}{p^j}\right)=(K-1)z^j=Kz^j-w_0^j\text{ for all } j.$$

Hence, the second inequality of (27) holds. Since $p_j\ge\sum_{i'=1}^{n}\alpha_{i'}^j+\frac{(1+v_0^j)\alpha_i^j}{v_i^j}$ for any i,j, we have

$$w_i^j=\frac{\alpha_i^j}{p^j v_i^j}\le\frac{1}{1+v_0^j}\left(1-\sum_{i'=1}^{n}\frac{\alpha_{i'}^j}{p^j}\right)=z^j\text{ for all } i\in\mathcal{N},j\in\mathcal{M}.$$

Therefore, (27) holds. Hence, the first-stage feasible region of $\boldsymbol{\alpha}$ is characterized by (32). This completes the proof of Proposition 1. □

Proof of Theorem 2

Let us consider a problem identical to $\mathcal{OP}(\gamma)$ but without budget constraints (i.e., $B_i(\gamma)=+\infty$ for all $i\in\mathcal{N}$ and $\gamma>0$), which we denote as $\mathcal{OP}_*(\gamma)$. By definition, in $\mathcal{OP}_*(\gamma)$, ad i will not run out of budget throughout the planning horizon. Throughout the proof of Theorem 2, we write $\tilde{y}_i^j(t)=\tilde{y}_i^j(t|\tilde{\pi}_{\mathsf{DWO}}(\boldsymbol{\alpha}),\mathcal{H}_{t-1})$ whenever there is no confusion.

- *Step 1.* For problem $\mathcal{OP}_*(\gamma)$, if $\boldsymbol{\alpha}$ is single-period feasible, it holds that

$$\liminf_{\gamma\uparrow+\infty}\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}\tilde{y}_i^j(t)\ge\alpha_i^j\text{ almost surely for all } i\in\bar{\mathcal{N}}\text{ and } j\in\mathcal{M}.\tag{33}$$

Under the DWO-$\boldsymbol{\alpha}$ algorithm, we have that

$$t\alpha_i^j-\sum_{\tau=1}^{t}\tilde{y}_i^j(t)=\tilde{d}_i^j(t+1)\le\left(\tilde{d}_i^j(t+1)\right)^+.$$

Therefore, it suffices to show that, if (11) holds,

$$\limsup_{t\to+\infty}\frac{1}{t}\cdot\left(\tilde{d}_i^j(t+1,s)\right)^+\le 0\text{ with probability 1 in problem }\mathcal{OP}_*(+\infty).$$

For a vector $\boldsymbol{x} \in \mathbb{R}^n$, we use $\boldsymbol{x}^+$ to denote the component-wise positive part of $\boldsymbol{x}$. Also note that, for any $A, B \in \mathbb{R}$, $((A+B)^+)^2 \leq (A^+ + B)^2$. We have

$$\begin{aligned}\mathbb{E}||\big(\tilde{\boldsymbol{d}}(t+1)\big)^+||_2^2 =& \mathbb{E}||\big(\tilde{\boldsymbol{d}}(t)+\boldsymbol{\alpha}-\tilde{\boldsymbol{y}}(t)\big)^+||_2^2 \leq \mathbb{E}||(\tilde{\boldsymbol{d}}(t))^+ + \boldsymbol{\alpha} - \tilde{\boldsymbol{y}}(t)||_2^2 \\ =& \mathbb{E}||(\tilde{\boldsymbol{d}}(t))^+||_2^2 + \mathbb{E}||\boldsymbol{\alpha}-\tilde{\boldsymbol{y}}(t)||_2^2 + 2\mathbb{E}\left[\sum_{i,j}(\tilde{d}_i^j(t))^+ \cdot \alpha_i^j - \sum_{i,j}(\tilde{d}_i^j(t))^+ \cdot \tilde{y}_i^j(t)\right],\end{aligned}$$

where $||\cdot||_2$ denotes the ℓ_2-norm in a Euclidean space. Since $(\tilde{d}_i^j(t))^+ \geq 0$ for all i and j, inequality (11) implies that

$$\mathbb{E}\left[\sum_{i,j}(\tilde{d}_i^j(t))^+ \cdot \alpha_i^j - \sum_{i,j}(\tilde{d}_i^j(t))^+ \cdot \tilde{y}_i^j(t)\right] \leq 0.$$

Furthermore,

$$\mathbb{E}||\boldsymbol{\alpha}-\tilde{\boldsymbol{y}}(t)||_2^2 \leq (n+1)\cdot m \cdot C, \text{ where } C := \max \mathbb{E}[\tilde{y}_i^j(t)] < +\infty.$$

Therefore,

$$\mathbb{E}||(\tilde{\boldsymbol{d}}(t+1))^+||_2^2 \leq ||(\tilde{\boldsymbol{d}}(1))^+||_2^2 + t(n+1)mC \text{ for all } t \geq 1.$$

By Jensen's inequality and that $||\cdot||_2^2$ is convex,

$$||\mathbb{E}[(\tilde{\boldsymbol{d}}(t+1))^+]||_2^2 \leq \mathbb{E}||(\tilde{\boldsymbol{d}}(t+1))^+||_2^2 \leq t(n+1)mC \text{ for all } t \geq 1.$$

Therefore,

$$0 \leq \frac{1}{t}||\mathbb{E}[(\tilde{\boldsymbol{d}}(t+1))^+]||_2 \leq \sqrt{\frac{(n+1)mC}{t}}, \text{ which implies that } \limsup_{t\to+\infty}\frac{1}{t}||\mathbb{E}[(\tilde{\boldsymbol{d}}(t+1))^+]||_2 = 0.$$

Hence

$$\limsup_{t\to+\infty}\frac{1}{t}\cdot(\tilde{d}_i^j(t))^+ = 0 \text{ with probability 1, for all } i \in \bar{\mathcal{N}} \text{ and } j \in \mathcal{M}.$$

Inequality (33) then follows immediately.

- *Step 2.* For problem $\mathcal{OP}_*(\gamma)$, if $\boldsymbol{\alpha}$ is single-period feasible, it holds that

$$\limsup_{\gamma\uparrow+\infty}\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}\tilde{y}_i^j(t) \leq \alpha_i^j \text{ almost surely for all } i \in \bar{\mathcal{N}} \text{ and } j \in \mathcal{M}. \tag{34}$$

Assume that, to the contrary, there exists (i_0, j_0) such that

$$\limsup_{\gamma\uparrow+\infty}\frac{\sum_{t=1}^{T(\gamma)}\tilde{y}_{i_0}^{j_0}(t)}{T(\gamma)} > \alpha_{i_0}^{j_0}.$$

Hence, there exists some $\Delta > 0$, such that

$$\frac{\sum_{t=1}^{T(\gamma)}\tilde{y}_{i_0}^{j_0}(t)}{T(\gamma)} > \alpha_{i_0}^{j_0} + \Delta \text{ for infinitely many } \gamma. \tag{35}$$

Denote the set of γ's that satisfy (35) as Γ. Note that $\frac{1}{T(\gamma)}(\sum_{t=1}^{s}\tilde{y}_{i_0}^{j_0}(t))$ increases by at most $1/(T(\gamma))$ as s increases by 1. Hence, for all $\gamma \in \Gamma$ and $\gamma > 3/(T\Delta)$, $\frac{1}{T(\gamma)}(\sum_{t=1}^{s}\tilde{y}_{i_0}^{j_0}(t))$ increases by no more than $\Delta/3$ if s increases by 1. Therefore, for all $\gamma \in \Gamma$ and $\gamma > 3/(T\Delta)$, there exists a $s(\gamma) < T(\gamma)$, such that

$$\alpha_{i_0}^{j_0} + \frac{\Delta}{3} < \frac{\sum_{t=1}^{s(\gamma)}\tilde{y}_{i_0}^{j_0}(t)}{T(\gamma)} < \alpha_{i_0}^{j_0} + \frac{2\Delta}{3} \tag{36}$$

By (36), we have that, for infinitely many γ,

$$\sum_{t=1}^{s(\gamma)} \tilde{y}_{i_0}^{j_0}(t) > T(\gamma)\left(\alpha_{i_0}^{j_0} + \frac{\Delta}{3}\right).$$

Hence, for infinitely many γ,

$$(\tilde{d}_{i_0}^{j_0}(t))^+ = \left((t-1)\alpha_{i_0}^{j_0} - \sum_{\tau=1}^{t-1} \tilde{y}_{i_0}^{j_0}(\tau)\right)^+ = 0 \text{ for all } t \geq s(\gamma)+1,$$

where the equality follows from

$$\sum_{\tau=1}^{t-1} \tilde{y}_{i_0}^{j_0}(\tau) \geq \sum_{\tau=1}^{s(\gamma)} \tilde{y}_{i_0}^{j_0}(\tau) > T(\gamma)\alpha_{i_0}^{j_0} > (t-1)\alpha_{i_0}^{j_0}.$$

Therefore, ad i_0 will not be offered to customer type j_0 for all $t \geq s(\gamma)+1$. Hence, $\tilde{y}_{i_0}^{j_0}(t) = 0$ for all $t \geq s(\gamma)+1$ and $t \leq T(\gamma)$. By (36), we have

$$\frac{\sum_{t=1}^{T(\gamma)} \tilde{y}_{i_0}^{j_0}(t)}{T(\gamma)} = \frac{\sum_{t=1}^{s(\gamma)} \tilde{y}_{i_0}^{j_0}(t)}{T(\gamma)} < \alpha_{i_0}^{j_0} + \frac{2\Delta}{3} \text{ for } \gamma \in \Gamma \text{ and } \gamma > \frac{3}{T\Delta},$$

which contradicts inequality (35). Therefore, for the system of $\mathcal{OP}_*(\gamma)$, we have inequality (34) holds.

- *Step 3.* For problem $\mathcal{OP}(\gamma)$, if $\boldsymbol{\alpha}$ is first-stage feasible, then Eq. (15) holds.

Inequalities (33) and (34) together imply that (15) holds for problem $\mathcal{OP}_*(\gamma)$. In particular, for problem $\mathcal{OP}_*(\gamma)$, we have no stock-out occurs for any ad asymptotically, i.e.,

$$\lim_{\gamma\uparrow+\infty} \frac{1}{T(\gamma)} \sum_{j\in\mathcal{M}} \sum_{t=1}^{T(\gamma)} \tilde{y}_i^j(t) = \alpha_i^j \leq \frac{B_i(\gamma)}{b_i T(\gamma)} = \frac{B_i}{b_i T} \text{ for all } i \in \bar{\mathcal{N}}. \tag{37}$$

Furthermore, by construction, the click-through process of $\mathcal{OP}_*(\gamma)$ is *identical* to that of $\mathcal{OP}(\gamma)$ before stock-out occurs in $\mathcal{OP}(\gamma)$, which occurs with probability 0 as $\gamma \uparrow +\infty$ by (37). Therefore, for the problem $\mathcal{OP}(\gamma)$, a standard coupling argument implies that (15) holds as well. This completes the proof of Theorem 2. □

Proof of Theorem 3

The theorem follows directly from inequalities (18) in Proposition 2 and *Steps 1 and 2* in the proof of Proposition 2. □

Proof of Lemma 1

We first check the feasibility of $\hat{\boldsymbol{\alpha}}(\boldsymbol{z})$ by directly plugging $\hat{\alpha}_i^j(\boldsymbol{z})$ into the constraints of ($\mathcal{OTP}$). Since $\boldsymbol{z}$ is feasible to $\mathcal{FA}(\gamma)$, we have, for all $i \in \mathcal{N}$,

$$b_i \sum_{j\in\mathcal{L}_i} \hat{\alpha}_i^j(\boldsymbol{z}) = b_i \sum_{j\in\mathcal{L}_i, i\in S\in\mathfrak{S}^j} p^j \phi_i^j(S) z^j(S) \leq \frac{B_i}{T};$$

and, for all $i \in \bar{\mathcal{N}}$ and $\mathcal{C} \in \mathfrak{K}_i$,

$$\sum_{j\in\mathcal{C}} \hat{\alpha}_i^j(\boldsymbol{z}) = b_i \sum_{j\in\mathcal{L}_i, i\in S\in\mathfrak{S}^j} p^j \phi_i^j(S) z^j(S) \geq \frac{\eta_i^{\mathcal{C}}}{T}.$$

In addition,

$$\mathcal{V}_{\mathsf{CT}}(\hat{\boldsymbol{\alpha}}(\boldsymbol{z})) = \sum_{i\in\bar{\mathcal{N}}} \sum_{j\in\mathcal{L}_i} \sum_{i\in S\in\mathfrak{S}^j} r_i^j p^j \phi_i^j(S) z^j(S) + \lambda F(\hat{\boldsymbol{\alpha}}(\boldsymbol{z})) = \sum_{i\in\bar{\mathcal{N}}, j\in\mathcal{M}, S\in\mathfrak{S}^j} r_i^j p^j \phi_i^j(S) z^j(S) + \lambda F(\hat{\boldsymbol{\alpha}}(\boldsymbol{z})) = \mathcal{V}_{\mathsf{FA}}(\boldsymbol{z}).$$

Therefore, it remains to show that

$$\mathbb{E}_{\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}),\tilde{j}(t),\mathsf{D}_y}[\tilde{y}_i^j(t|\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}))] = \hat{\alpha}_i^j(\boldsymbol{z}). \tag{38}$$

Applying the law of total probability, we directly evaluate that

$$\begin{aligned}\mathbb{E}_{\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}),\tilde{j}(t),\mathsf{D}_y}[\tilde{y}_i^j(t|\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}))] =& \mathbb{P}[\tilde{j}(t)=j]\sum_{S\in\mathfrak{S}^j}\mathbb{P}[\tilde{\pi}_{\mathsf{AF}}(j)=S]\mathbb{E}[\tilde{y}_i^j(t)|\tilde{\pi}_{\mathsf{AF}}(j)=S,\tilde{j}(t)=j]\\ =& p^j\sum_{i\in S\in\mathfrak{S}^j} z^j(S)\phi_i^j(S)\\ =& \hat{\alpha}_i^j(\boldsymbol{z}),\end{aligned}$$

i.e., (38) holds. Therefore, the click-through target vector $\hat{\boldsymbol{\alpha}}(\boldsymbol{z})$ is first-stage feasible. In particular, $\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)$ is first-stage feasible with $\mathbb{E}_{\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*),\tilde{j}(t),\mathsf{D}_y}[\tilde{y}_i^j(t|\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*))] = \hat{\alpha}_i^j(\boldsymbol{z}^*)$. We defer the proof of $\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)$'s optimality for $(\mathcal{OTP})$ to the proof of Proposition 2. □

Proof of Proposition 2

We prove (18) by showing each individual inequality thereof.

- *Step 1.* The optimal FAV of $(\mathcal{OTP})$ dominates that of $\mathcal{FA}(\gamma)$, i.e.,

$$\mathcal{V}^*_{\mathsf{CT}} \geq \mathcal{V}^*_{\mathsf{FA}}. \tag{39}$$

As shown by Lemma 1, $\hat{\boldsymbol{\alpha}}(\boldsymbol{z})$ is feasible in $(\mathcal{OTP})$ and, thus,

$$\mathcal{V}^*_{\mathsf{CT}} \geq \mathcal{V}_{\mathsf{CT}}(\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)) = \mathcal{V}_{\mathsf{FA}}(\boldsymbol{z}^*) = \mathcal{V}^*_{\mathsf{FA}}. \tag{40}$$

Inequality (39) thus follows from (40).

- *Step 2.* The FAV generated by $\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*)$ in $\mathcal{OP}(\gamma)$ is asymptotically identical to the optimal FAV in $\mathcal{FA}(\gamma)$, i.e.,

$$\lim_{\gamma\uparrow}\mathcal{V}(\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*)|\gamma) = \mathcal{V}^*_{\mathsf{FA}}. \tag{41}$$

We first show that $\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*)$ is asymptotically *feasible* for $\mathcal{OP}(\gamma)$. As $\gamma\uparrow+\infty$, we have

$$\lim_{\gamma\uparrow+\infty}\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}\sum_{j\in\mathcal{M}} b_i\tilde{y}_i^j(t|\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*)) = \mathbb{E}\left[\sum_{j\in\mathcal{M}} b_i\tilde{y}_i^j(t|\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*))\right] = \sum_{j\in\mathcal{M},S\in\mathfrak{S}^j} b_i p^j\phi_i^j(S)z^{j*}(S) \leq \frac{B_i(\gamma)}{T(\gamma)}, \tag{42}$$

where the first equality follows from the strong law of large numbers, the second from the definition of $\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*)$, and the inequality follows from $\boldsymbol{z}^*$ is feasible for $\mathcal{FA}(\gamma)$. Similarly, we have, under the policy $\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*)$,

$$\lim_{\gamma\uparrow+\infty}\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}\sum_{j\in\mathcal{C}}\tilde{y}_i^j(t|\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*)) = \mathbb{E}\left[\sum_{j\in\mathcal{C}}\tilde{y}_i^j(t|\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*))\right] = \sum_{j\in\mathcal{C},S\in\mathfrak{S}^j} p^j\phi_i^j(S)z^j(S) \geq \frac{\eta_i^{\mathcal{C}}(\gamma)}{T(\gamma)}, \tag{43}$$

where the first equality follows from the strong law of large numbers, the second from the definition of $\tilde{\pi}_{\mathsf{FA}}(\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*))$, and the inequality follows from $\boldsymbol{z}^*$ is feasible for $\mathcal{FA}(\gamma)$. Inequalities (42) and (43) together imply that $\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*)$ is asymptotically feasible for $\mathcal{OP}(\gamma)$ as $\gamma\uparrow+\infty$.

Then, we evaluate the FAV of the FA-$\boldsymbol{z}^*$ policy as follows:

$$\begin{aligned}\lim_{\gamma\uparrow+\infty}\mathcal{V}(\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*)|\gamma) =& \lim_{\gamma\uparrow+\infty}\mathbb{E}\left[\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}\sum_{i\in\bar{\mathcal{N}}}\sum_{j\in\mathcal{M}} r_i^j\tilde{y}_i^j(t|\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*)) + \lambda F(\bar{\boldsymbol{y}}(\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*)))\right]\\ =& \sum_{i\in\bar{\mathcal{N}},j\in\mathcal{M},S\in\mathfrak{S}^j} r_i^j p^j\phi_i^j(S)z^{j*}(S) + \lambda F(\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*))\\ =& \mathcal{V}_{\mathsf{FA}}(\boldsymbol{z}^*)\\ =& \mathcal{V}^*_{\mathsf{FA}},\end{aligned} \tag{44}$$

where the second equality follows from the law of large numbers and the dominated convergence theorem. Therefore, (41) follows from (44).

- *Step 3.* The optimal FAV of $\mathcal{FA}(\gamma)$ dominates that of $\mathcal{OP}(\gamma)$, i.e.,

$$\mathcal{V}^*_{\mathsf{FA}} \geq \mathcal{V}^*(\gamma) \text{ for any } \gamma > 0. \tag{45}$$

Consider an arbitrary policy $\tilde{\pi} \in \tilde{\Pi}$ feasible for $\mathcal{OP}(\gamma)$. We first define the following probability measure induced by $\tilde{\pi}$, $\boldsymbol{z}_{\mathsf{FA}}(\tilde{\pi})$ for $\mathcal{FA}(\gamma)$: For $j \in \mathcal{M}$ and $S \in \mathfrak{S}^j$,

$$z^j_{\mathsf{FA}}(S|\tilde{\pi}) := \mathbb{P}\left[\tilde{\pi}_{\tilde{t}}(\tilde{j}(\tilde{t}), \mathcal{H}_{\tilde{t}-1}) = S | \tilde{j}(\tilde{t}) = j\right], \tag{46}$$

where $\tilde{t}$ is a random variable uniformly distributed on $\{1, 2, ..., T(\gamma)\}$ and independent everything else. Because $\tilde{\pi}$ is feasible for $\mathcal{OP}(\gamma)$, all the constraints of $\mathcal{OP}(\gamma)$ will also be satisfied in the expected sense as well, i.e.,

$$\frac{1}{T(\gamma)}\mathbb{E}\left[\sum_{t=1}^{T(\gamma)}\sum_{j\in\mathcal{M}} b_i \tilde{y}^j_i(t|\tilde{\pi}, \mathcal{H}_{t-1})\right] \leq \frac{B_i(\gamma)}{T(\gamma)}, \text{ for each } i \in \mathcal{N},$$

and

$$\frac{1}{T(\gamma)}\mathbb{E}\left[\sum_{t=1}^{T}\sum_{j\in\mathcal{C}} \tilde{y}^j_i(t|\tilde{\pi}, \mathcal{H}_{t-1})\right] \geq \frac{\eta^{\mathcal{C}}_i}{T(\gamma)}, \quad \text{for each } i \in \bar{\mathcal{N}} \text{ and } \mathcal{C} \in \mathfrak{K}_i,$$

where the expectations are taken with respect to $\tilde{j}(t)$, $\tilde{\pi}$, and $\tilde{\boldsymbol{y}}$. By (46), straightforward algebraic manipulation and the law of iterated expectations together yield that

$$\mathbb{E}[\bar{y}^j_i] = \mathbb{E}\left[\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)} \tilde{y}^j_i(t|\tilde{\pi}_t, \mathcal{H}_{t-1})\right] = \sum_{j\in\mathcal{M}, S\in\mathfrak{S}^j} p^j \phi^j_i(S) z^j_{\mathsf{FA}}(S|\tilde{\pi}).$$

Plugging this identity into the constraints of $\mathcal{FA}(\gamma)$, we have

$$\sum_{j\in\mathcal{M}, S\in\mathfrak{S}^j} b_i p^j \phi^j_i(S) z^j_{\mathsf{FA}}(S|\tilde{\pi}) \leq \frac{B_i(\gamma)}{T(\gamma)} \text{ for all } i \in \mathcal{N}$$

and that

$$\sum_{j\in\mathcal{C}, S\in\mathfrak{S}^j} p^j \phi^j_i(S) z^j_{\mathsf{FA}}(S|\tilde{\pi}) \geq \frac{\eta^{\mathcal{C}}_i(\gamma)}{T(\gamma)} \text{ for all } i \in \bar{\mathcal{N}} \text{ and } \mathcal{C} \in \mathfrak{K}_i.$$

Therefore, $\boldsymbol{z}_{\mathsf{FA}}(\tilde{\pi})$ is feasible for $\mathcal{FA}(\gamma)$ and, hence,

$$\mathcal{V}^*_{\mathsf{FA}} \geq \mathcal{V}_{\mathsf{FA}}(\boldsymbol{z}_{\mathsf{FA}}(\tilde{\pi})). \tag{47}$$

By Jensen's inequality and the concavity of the fairness metric $F(\cdot)$,

$$\begin{aligned} \mathcal{V}_{\mathsf{FA}}(\boldsymbol{z}_{\mathsf{FA}}(\tilde{\pi})) &= \sum_{i\in\bar{\mathcal{N}}}\sum_{j\in\mathcal{L}_i}\sum_{i\in S\in\mathfrak{S}^j} r^j_i p^j \phi^j_i(S) z^j_{\mathsf{FA}}(S|\tilde{\pi}) + \lambda F(\hat{\boldsymbol{\alpha}}(\boldsymbol{z}_{\mathsf{FA}}(S|\tilde{\pi}))) \\ &\geq \mathbb{E}\left[\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}\sum_{i\in\bar{\mathcal{N}}}\sum_{j\in\mathcal{M}} r^j_i \tilde{y}^j_i(t|\tilde{\pi}_t, \mathcal{H}_{t-1}) + \lambda F(\bar{\boldsymbol{y}}(\tilde{\pi}))\right] \\ &= \mathcal{V}(\tilde{\pi}|\gamma) \end{aligned} \tag{48}$$

Since $\tilde{\pi}$ is arbitrary, inequalities (47) and (48) together imply that (45) holds.

- *Step 4.* The following inequality holds:

$$\mathcal{V}^*_{\mathsf{FA}} = \lim_{\gamma\uparrow+\infty} \mathcal{V}(\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*)|\gamma) = \lim_{\gamma\uparrow+\infty} \mathcal{V}^*(\gamma) \geq \lim_{\gamma\uparrow+\infty} \mathcal{V}(\tilde{\pi}_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma) \tag{49}$$

Inequality (45), equality (41), and the trivial inequality $\mathcal{V}^*(\gamma) \geq \mathcal{V}(\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*)|\gamma)$ for any $\gamma > 0$ together imply that

$$\mathcal{V}^*_{\mathsf{FA}} = \lim_{\gamma\uparrow+\infty} \mathcal{V}(\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*)|\gamma) = \lim_{\gamma\uparrow+\infty} \mathcal{V}^*(\gamma).$$

The optimality of $\mathcal{V}^*(\gamma)$ implies that $V^*(\gamma) \geq \mathcal{V}(\tilde{\pi}_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma)$ for all $\gamma > 0$. Putting everything together, we have the inequality (49) follows.

- *Step 5.* The DWO-$\boldsymbol{\alpha}^*$ policy generates the same asymptotic FAV in $\mathcal{OP}$ as the optimal FAV in $(\mathcal{OTP})$, i.e.,

$$\lim_{\gamma\uparrow+\infty} \mathcal{V}(\tilde{\pi}_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma) = \mathcal{V}^*_{\mathsf{CT}} \tag{50}$$

By Theorem 2, the DWO-$\boldsymbol{\alpha}^*$ policy is asymptotically feasible for the original problem $(\mathcal{OP})$ with (15) holding true. We now evaluate the FAV $\mathcal{V}(\tilde{\pi}_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma)$ in the asymptotic regime:

$$\begin{aligned}
\lim_{\gamma\uparrow+\infty} \mathcal{V}(\tilde{\pi}_{\mathsf{DOW}}(\boldsymbol{\alpha}^*)|\gamma) = &\lim_{\gamma\uparrow+\infty} \mathbb{E}\left[\frac{1}{T(\gamma)} \sum_{t=1}^{T(\gamma)} \sum_{i\in\bar{\mathcal{N}}} \sum_{j\in\mathcal{M}} r_i^j \tilde{y}_i^j(t|\tilde{\pi}_{\mathsf{DOW}}(\boldsymbol{\alpha}^*), \mathcal{H}_{t-1}) + \lambda F(\bar{\boldsymbol{y}}(\tilde{\pi}_{\mathsf{DOW}}(\boldsymbol{\alpha}^*)))\right] \\
= &\sum_{i\in\bar{\mathcal{N}}, j\in\mathcal{M}} r_i^j \alpha_i^{j*} + \lambda F(\boldsymbol{\alpha}^*) \\
= &\mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}^*) \\
= &\mathcal{V}^*_{\mathsf{CT}},
\end{aligned} \tag{51}$$

where the second equality follows from equality (15) and the dominated convergence theorem. Hence, equality (50) follows from equality (51), i.e.,

Therefore, putting the (in)equalities (39), (49), and (50) together, we have (18) holds, which completes the proof of Proposition 2. As a by product of the proof, (18) also implies that the inequality in (40) holds as equality, i.e., $\mathcal{V}^*_{\mathsf{CT}} = \mathcal{V}_{\mathsf{CT}}(\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*))$. Hence, $\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)$ is optimal for $(\mathcal{OTP})$, which also completes the proof of Lemma 1. □

Proof of Proposition 3

The proof follows from the same argument as *Step 5* in the proof of Proposition 2 by replacing $\boldsymbol{\alpha}^*$ with any feasible $\boldsymbol{\alpha}$. To avoid repetition, we omit the proof details. □

E. Feasible Click-Through Targets Under the MNL Choice Model

Proposition 1 characterizes the feasible region of the click-through targets $\mathcal{A}_{MNL}$ if customers follow the MNL choice model. This section seeks to deliver additional insights on when the click-through targets are feasible. We observe that (32) is equivalent to

$$p^j \geq \sum_{i=1}^{n} \alpha_i^j + \max\left\{\frac{1+v_0^j}{K-1} \sum_{i=1}^{n} \frac{\alpha_i^j}{v_i^j}, \max_i \left\{\frac{(1+v_0^j)\alpha_i^j}{v_i^j}\right\}\right\}, \text{ for all } j.$$

Here, p^j is the expected (per-user) traffic of type j customers in each period. Clearly, $\sum_{i=0}^{n} \alpha_i^j$ is the total required traffic for type j customers if a customer will click one of the ad in the offer set with probability 1.

In practice, however, a customer may end up not choosing any ad from the offer set, so we need some buffer traffic for type-j customers that accounts for the nonclick circumstance.

More specifically, let $\mathfrak{S}_i$ denote the collection of all offer sets containing ad i. Since the offer-set policy may be random, we define $\mu_j(S)$ as the probability of displaying offer-set $S \subseteq \bar{\mathcal{N}}$ to type j customers. Thus, the desired click-through goal for ad i and customer-type j customer is

$$\sum_{S\in\mathfrak{S}_i} \mu_j(S)\cdot \frac{v_i^j}{1+\sum\limits_{i'\in S} v_{i'}^j} \geq \alpha_i^j.$$

Thus, the nonclick probability of the ads for a type-j customer when ad i ($i\in\bar{\mathcal{N}}$) is offered satisfies that

$$\alpha_i^j(o) := \sum_{S\in\mathfrak{S}_i} \mu_j(S)\cdot \frac{1+v_0^j}{1+\sum_{i'\in S} v_{i'}^j} \geq \frac{(1+v_0^j)\alpha_i^j}{v_i^j}$$

Therefore, to ensure the click-through goal of type j customers and ad i, the traffic of customer type j must satisfy $p^j \geq \sum_{i'} \alpha_{i'}^j + \alpha_i^j(o) \geq \sum_{i'} \alpha_{i'}^j + \frac{(1+v_0^j)\alpha_i^j}{v_i^j}$ for all $i\in\bar{\mathcal{N}}$.

The cardinality constraint for the offer set size would impose an additional bound on the nonpurchase probability of type j customers. Specifically, let $\mathfrak{S} := \bigcup_{i=1}^n \mathfrak{S}_i$ be the set of all offer sets displayed to a customer. Because $|S|\leq K$ for any $S\in\mathfrak{S}$, $|\{i\in\bar{\mathcal{N}} : S\in\mathfrak{S}_i\}|\leq K$ for all S. Note that ad 0 is always included in the offer set. We have, given customer type j,

$$(K-1)\sum_{S\in\mathfrak{S}} \mu_j(S)\cdot \frac{1+v_0^j}{1+\sum_{i\in S} v_i^j} \geq \sum_{i=1}^n \sum_{S\in\mathfrak{S}_i} \mu_j(S)\cdot \frac{1+v_0^j}{1+\sum_{i\in S} v_i^j} \geq \sum_{i=1}^n \frac{(1+v_0^j)\alpha_i^j}{v_i^j}.$$

Thus, the nonpurchase probability of all ads for type-j customer satisfies that

$$\alpha_o^j := \sum_{S\in\mathfrak{S}} \mu_j(S)\cdot \frac{1+v_0^j}{1+\sum_{i\in S} v_i^j} \geq \frac{1+v_0^j}{K-1}\sum_{i=1}^n \frac{\alpha_i^j}{v_i^j}.$$

Therefore, given the cardinality constraint of an offer set, to ensure the click-through goals of type-j customers with respect to all ads, the traffic of customer-type j must satisfy $p^j \geq \sum_{i=1}^n \alpha_i^j + \alpha_o^j \geq \sum_i \alpha_i^j + \frac{1+v_0^j}{K-1}\sum_{i=1}^n \frac{\alpha_i^j}{v_i^j}$. In summary, the characterization for the feasibility of $\boldsymbol{\alpha}$ demonstrates that, to meet the click-through targets, we should also account for the *nonclick* cases.

F. Optimal Target Convex Program Formulation for Specific Choice Models

In this section, we introduce the characterization of first-stage feasible region $\mathcal{A} := \{\boldsymbol{\alpha}\in[0,1]^{(n+1)m} : h(\boldsymbol{\alpha}) \geq 0\}$ for independent and generalized attraction choice models. Similar to the characterization of $\mathcal{A}$ if the customers follow the MNL model, we use a binary variable representation of a deterministic stationary history-independent offer-set policy $\pi\in\Pi_{\mathsf{s}}$ (see, also, the proof of Proposition 1). More specifically, $\pi\in\Pi_{\mathsf{s}}$ can be equivalently represented by an $(n+1)m-$dimensional binary vector $\boldsymbol{x} = (x_i^j\in\{0,1\} : i\in\bar{\mathcal{N}}, j\in\mathcal{M})$, where $x_i^j = 1$ means that $i\in\pi(j)$, i.e., ad i is included in the offer-set displayed to a type j customer. Because the organic recommendation is always included in an offer-set, we have $x_0^j = 1$ for all $j\in\mathcal{M}$. We denote $\phi_i^j(\boldsymbol{x})$ as the expected click-throughs of a type j customer for ad i if the offer-set displayed to this customer is $S^j = \{i\in\bar{\mathcal{N}} : x_i^j = 1\}$. Denote the set of all plausible offer-set representation vectors as $\mathcal{X}\subset\{0,1\}^{(n+1)m}$, and the set of plausible offer-set representation vectors displayed to a type j customer as $\mathcal{X}^j\subset\{0,1\}^{(n+1)}$.

F.1. Independent Choice Model

If customers follow the independent choice model, the click-throughs only depend on the customer type j and ad i, but not on the offer-set S^j displayed to the customer. This is actually the most commonly adopted choice models in practical advertising (see, e.g., Feldman et al. 2021), where the click-through rate (CTR) prediction algorithm of the platform outputs:

$$\phi_i^j(\boldsymbol{x}) = c_i^j x_i^j, \tag{52}$$

where $c_i^j > 0$ is the CTR of ad i with respect to a type-j customer. It is evident from (52) that the number of click-throughs is independent of the ads in S^j other than ad i. The following proposition is a counterpart of Proposition 1 with the independent choice model.

Proposition 4. *If customers follow the independent click-through model* (52)*, the first-stage feasible region $\mathcal{A}_{IND}$ is given by the following linear constraints:*

$$\mathcal{A}_{IND} = \left\{\boldsymbol{\alpha} \in [0,1]^{(n+1)m} : p^j c_i^j \geq \alpha_i^j, \forall i \in \bar{\mathcal{N}}, j \in \mathcal{L}_i\right\}. \tag{53}$$

It is expected that, under the independent choice model, we have an independent feasibility condition for each ad-customer type pair. With Proposition 4, we we can simplify the optimal target problem $(\mathcal{OTP})$ under the independent choice model as the following convex program with linear constraints only:

$$\begin{aligned}
\max_{\boldsymbol{\alpha} \geq \mathbf{0}} \ & \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}) \\
\text{s.t. } & p^j c_i^j \geq \alpha_i^j, \ \text{for each } i \in \bar{\mathcal{N}}, j \in \mathcal{L}_i \\
& b_i \sum_{j \in \mathcal{L}_i} \alpha_i^j \leq \frac{B_i}{T}, \ \text{for each } i \in \mathcal{N}, \\
& \sum_{j \in \mathcal{C}} \alpha_i^j \geq \frac{\eta_i^{\mathcal{C}}}{T}, \quad \text{for each } i \in \bar{\mathcal{N}} \text{ and } \mathcal{C} \in \mathfrak{K}_i.
\end{aligned} \tag{$\mathcal{OTP}-\mathcal{IND}$}$$

F.2. Generalized Attraction Model

The generalized attraction model (GAM) is a generalization of MNL accounts for the possibility that a customer may look for a product outside the offer-set (see, e.g., Gallego et al. 2015). Under the GAM, the expected number of click-throughs of ad i by a type-j customer is given by

$$\phi_i^j(\boldsymbol{x}) = \frac{v_i^j x_i^j}{1 + \sum_{i' \in \bar{\mathcal{N}}} \omega_{i'}^j (1 - x_{i'}^j) + \sum_{i \in \bar{\mathcal{N}}} v_{i'}^j x_{i'}^j}, \tag{54}$$

where $v_i^j > 0$ is the attraction value of ad i to a type-j customer, and $w_i^j \in [0, v_i^j]$ is the shadow attraction value of ad i to a type-j customer, capturing the customer's looking for a product outside the offer-set. Hence, by defining $\tilde{v}^j := 1 + \sum_{i \in \bar{\mathcal{N}}} \omega_i^j > 0$ and $\tilde{v}_i^j := v_i^j - \omega_i^j \geq 0$, we have, under the GAM,

$$\phi_i^j(\boldsymbol{x}) = \frac{v_i^j x_i^j}{\tilde{v}^j + \sum_{i' \in \bar{\mathcal{N}}} \tilde{v}_{i'}^j x_{i'}^j}. \tag{55}$$

As Corollary 1, we first rewrite the necessary and sufficient condition for $\boldsymbol{\alpha}$ under the GAM.

Corollary 2. *If customers follow the GAM click-through model (55), a click-through target vector $\boldsymbol{\alpha}$ is single-period feasible if and only if, for each $j \in \mathcal{M}$*

$$\max_{\boldsymbol{x}^j \in \mathcal{X}^j} \sum_{i \in \bar{\mathcal{N}}} \frac{p^j v_i^j \theta_i^j x_i^j}{\tilde{v}^j + \sum_{i' \in \bar{\mathcal{N}}} \tilde{v}_{i'}^j x_{i'}^j} \geq \sum_{i \in \bar{\mathcal{N}}} \alpha_i^j \theta_i^j \ \textit{for any } \theta_i^j \geq 0 \ (i \in \bar{\mathcal{N}}). \tag{56}$$

Similar to Proposition 1, we can characterize the first-stage feasible region of the click-through target vector under the GAM, $\mathcal{A}_{GAM} := \{\boldsymbol{\alpha} : (56) \text{ holds for each } j.\}$, using linear constraints. The following proposition characterizes $\mathcal{A}_{GAM}$ and accounts for the offer-set cardinality constraint.

Proposition 5. *If customers follow the GAM (55) and the size of an offer-set cannot exceed K, the first-stage feasible region $\mathcal{A}_{GAM}$ is given by the following linear constraints:*

$$\mathcal{A}_{GAM} := \left\{ \boldsymbol{\alpha} \in \mathbb{R}_+^{(n+1)m} : \sum_{i'=1}^n \frac{\tilde{v}_{i'}^j \alpha_{i'}^j}{v_{i'}^j} + \frac{(\tilde{v}^j + \tilde{v}_0^j)\alpha_i^j}{v_i^j} \leq p^j, \forall i, j, \ \textit{and} \ \sum_{i=1}^n \frac{\tilde{v}_i^j \alpha_i^j}{v_i^j} + \frac{\tilde{v}^j + \tilde{v}_0^j}{K-1} \sum_{i=1}^n \frac{\alpha_i^j}{v_i^j} \leq p^j, \forall j \right\}. \tag{57}$$

With Proposition 5, we we can simplify the optimal target problem $(\mathcal{OTP})$ under the generalized attraction model as the following convex program with linear constraints only:

$$\begin{aligned}
&\max_{\boldsymbol{\alpha} \geq \mathbf{0}} \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}) \\
&\text{s.t.} \ \sum_{i'=1}^n \frac{\tilde{v}_{i'}^j \alpha_{i'}^j}{v_{i'}^j} + \frac{(\tilde{v}^j + \tilde{v}_0^j)\alpha_i^j}{v_i^j} \leq p^j, \ \text{for each } i \in \bar{\mathcal{N}}, j \in \mathcal{M}, \\
&\quad \sum_{i=1}^n \frac{\tilde{v}_i^j \alpha_i^j}{v_i^j} + \frac{\tilde{v}^j + \tilde{v}_0^j}{K-1} \sum_{i=1}^n \frac{\alpha_i^j}{v_i^j} \leq p^j, \ \text{for each } j \\
&\quad b_i \sum_{j \in \mathcal{L}_i} \alpha_i^j \leq \frac{B_i}{T}, \ \text{for each } i \in \mathcal{N}, \\
&\quad \sum_{j \in \mathcal{C}} \alpha_i^j \geq \frac{\eta_i^{\mathcal{C}}}{T}, \quad \text{for each } i \in \bar{\mathcal{N}} \text{ and } \mathcal{C} \in \mathfrak{K}_i.
\end{aligned} \tag{$\mathcal{OTP}-\mathcal{GAM}$}$$

F.3. Proofs

In this subsection, we give proofs of the technical results presented in Appendix F.

Proof of Proposition 4

Directly applying Theorem 1 to the independent choice model implies that $\boldsymbol{\alpha}$ is feasible if and only if the following inequality holds.

$$\max_{\boldsymbol{x} \in \mathcal{X}} \sum_{i \in \bar{\mathcal{N}}, j \in \mathcal{M}} p^j c_i^j \theta_i^j x_i^j \geq \sum_{i \in \bar{\mathcal{N}}, j \in \mathcal{M}} \alpha_i^j \theta_i^j \ \text{for any } \boldsymbol{\theta} \geq \mathbf{0}. \tag{58}$$

Setting $\theta_i^j = 1$ and all other θ's equal to zero in (58) immediately implies that if $\boldsymbol{\alpha}$ is single-period feasible, then (53) holds. Reversely, if (53) holds, then for any $\boldsymbol{\theta} \geq \mathbf{0}$, we have

$$\max_{\boldsymbol{x} \in \mathcal{X}} \sum_{i \in \bar{\mathcal{N}}, j \in \mathcal{M}} p^j c_i^j \theta_i^j x_i^j = \sum_{i \in \bar{\mathcal{N}}, j \in \mathcal{M}} p^j c_i^j \theta_i^j \geq \sum_{i \in \bar{\mathcal{N}}, j \in \mathcal{M}} \alpha_i^j \theta_i^j,$$

i.e., (58) holds and, therefore, $\boldsymbol{\alpha}$ is single-period feasible. This concludes the proof of Proposition 4. □

Proof of Corollary 2

The proof follows from exactly the same argument as that of Corollary 1. We omit the details to avoid repetition. □

Proof of Proposition 5

As in the proof of Proposition 1, we first state and prove the following auxiliary lemma.

LEMMA 3. *If customers follow the GAM click-through model (55) and the size of an offer-set cannot exceed K, we have $\boldsymbol{\alpha}$ is single-period feasible if and only if there exist $\boldsymbol{w} := (w_i^j : i \in \bar{\mathcal{N}}, j \in \mathcal{M})$ and $\boldsymbol{z} := (z^j : j \in \mathcal{M})$ that satisfy the following linear constraints*

$$\begin{aligned} &p^j v_i^j w_i^j \geq \alpha_i^j,\ w_i^j \leq z^j,\ 0 \leq w_i^j \leq 1 \quad \forall i, j \\ &\sum_{i=0}^n \tilde{v}_i^j w_i^j + \tilde{v}^j z^j = 1,\ \sum_{i=0}^n w_i^j \leq K z^j,\ w_0^j = z^j \quad \forall j, \end{aligned} \tag{59}$$

where $z^j := 1/(\tilde{v}^j + \sum_{i'} \tilde{v}_{i'}^j x_{i'}^j)$ and $w_i^j := x_i^j z^j = x_i^j/(\tilde{v}^j + \sum_{i'} \tilde{v}_{i'}^j x_{i'}^j)$.

Proof of Lemma 3 It is straightforward to check that the left-hand side of (56) is quasi-convex in $\boldsymbol{x}^j$ for all j, so there always exists a maximizer on the boundary of the feasible region. Thus, we can relax the binary constraint $x_i^j \in \{0,1\}$ to $x_i^j \in [0,1]$ in (56), which is therefore equivalent to

$$\max_{\boldsymbol{x}^j \in [0,1]^{(n+1)}, \sum_i x_i^j \leq K} \sum_{i \in \bar{\mathcal{N}}} \frac{p^j v_i^j \theta_i^j x_i^j}{\tilde{v}^j + \sum\limits_{i' \in \bar{\mathcal{N}}} \tilde{v}_{i'}^j x_{i'}^j} \geq \sum_{i \in \bar{\mathcal{N}}} \alpha_i^j \theta_i^j \text{ for all } \boldsymbol{\theta}^j \geq \boldsymbol{0} \text{ and } j \in \mathcal{M}. \tag{60}$$

We change the decision variable and define, for all $j \in \mathcal{M}$,

$$z^j := \frac{1}{\tilde{v}^j + \sum\limits_{i'} \tilde{v}_{i'}^j x_{i'}^j} \text{ and } w_i^j := x_i^j z^j = \frac{x_i^j}{\tilde{v}^j + \sum\limits_{i'} \tilde{v}_{i'}^j x_{i'}^j}.$$

Since $x_0^j = 1$ for all j, we have $w_0^j = z^j$ for all $j \in \mathcal{M}$. Then, we can rewrite (60) as, for any j,

$$\begin{aligned} &\min_{\boldsymbol{\theta} \geq \boldsymbol{0}} \left(\max_{w_i^j, z^j} \sum_{i \in \bar{\mathcal{N}}} p^j v_i^j w_i^j \theta_i^j - \sum_{i \in \bar{\mathcal{N}}} \alpha_i^j \theta_i^j \right) \geq 0 \\ &\text{s.t. } \sum_{i=0}^n \tilde{v}_i^j w_i^j + \tilde{v}^j z^j = 1, \\ &\qquad \sum_{i \in \bar{\mathcal{N}}} w_i^j \leq K z^j, \\ &\qquad w_i^j \leq z^j \text{ for all } i \in \bar{\mathcal{N}}, \\ &\qquad w_0^j = z^j. \end{aligned} \tag{61}$$

By Sion's minimax theorem, we can exchange the maximization and minimization operators so that (61) is equivalent to, for any $j \in \mathcal{M}$:

$$\begin{aligned} \max_{\boldsymbol{w}^j, z^j} \min_{\boldsymbol{\theta}^j \geq 0} \ &\sum_{i=0}^n \theta_i^j (p^j v_i^j w_i^j - \alpha_i^j) \geq 0, \\ \text{s.t. } &\sum_{i=0}^n \tilde{v}_i^j w_i^j + \tilde{v}^j z^j = 1, \\ &\sum_{i=0}^n w_i^j \leq K z^j, \\ &0 \leq w_i^j \leq z^j,\ \text{for all } i \in \bar{\mathcal{N}}, \\ &w_0^j = z^j. \end{aligned} \tag{62}$$

Therefore, (62) holds if and only if there exist $\boldsymbol{w}^j$ and z^j such that all the constraints in (62) hold and $\sum_{i=0}^{n} \theta_i^j (p^j v_i^j w_i^j - \alpha_i^j) \geq 0$ holds for all $\boldsymbol{\theta}^j \geq \mathbf{0}$, which is equivalent to $p^j v_i^j w_i^j - \alpha_i^j \geq 0$ for all $i \in \bar{\mathcal{N}}$. Therefore, (62) is equivalent to that, for any $j \in \mathcal{M}$,

$$\begin{aligned} & p^j v_i^j w_i^j - \alpha_i^j \geq 0, \\ & \sum_{i=0}^{n} \tilde{v}_i^j w_i^j + \tilde{v}^j z^j = 1, \\ & \sum_{i=0}^{n} w_i^j \leq K z^j, \\ & 0 \leq w_i^j \leq z^j, \text{ for all } i \in \bar{\mathcal{N}}, \\ & w_0^j = z^j. \end{aligned} \tag{63}$$

That (63) holds for all $j \in \mathcal{M}$ is equivalent to that (59) holds. This completes the proof of Lemma 3. □

We now prove Proposition 5 itself. We first show that if (59) holds, then $\boldsymbol{\alpha} \in \mathcal{A}_{GAM}$. By the first inequality of (59), we have $w_i^j \geq \frac{\alpha_i^j}{p^j v_i^j}$ for all i and j. Plugging this into the first and second equalities of (59), we have

$$1 - \tilde{v}^j z^j - \tilde{v}_0^j z^j = \sum_{i=1}^{n} \tilde{v}_i^j w_i^j \geq \sum_{i=1}^{n} \frac{\tilde{v}_i^j \alpha_i^j}{p^j v_i^j}.$$

Thus, by the first and second inequalities of (59), we have

$$\sum_{i'=1}^{n} \frac{\tilde{v}_{i'}^j \alpha_{i'}^j}{p^j v_{i'}^j} \leq 1 - \tilde{v}^j z^j - \tilde{v}_0^j z^j \leq 1 - (\tilde{v}^j + \tilde{v}_0^j) w_i^j \leq 1 - \frac{(\tilde{v}^j + \tilde{v}_0^j)\alpha_i^j}{p^j v_i^j} \text{ for any } i, j.$$

Rearranging the terms, we have

$$p^j \geq \sum_{i'=1}^{n} \frac{\tilde{v}_{i'}^j \alpha_{i'}^j}{v_{i'}^j} + \frac{(\tilde{v}^j + \tilde{v}_0^j)\alpha_i^j}{v_i^j} \text{ for any } i, j.$$

The first, second, and fourth inequalities of (59) imply that

$$\sum_{i=1}^{n} \frac{\alpha_i^j}{p^j v_i^j} \leq \sum_{i=1}^{n} w_i^j \leq (K-1) z^j = (K-1) \frac{1 - \sum_{i=1}^{n} \tilde{v}_i^j w_i^j}{\tilde{v}^j + \tilde{v}_0^j} \leq \frac{K-1}{\tilde{v}^j + \tilde{v}_0^j} \left(1 - \sum_{i=1}^{n} \frac{\tilde{v}_i^j \alpha_i^j}{p^j v_i^j}\right) \text{ for all } j.$$

Rearranging the terms, we have

$$p^j \geq \sum_{i=1}^{n} \frac{\tilde{v}_i^j \alpha_i^j}{v_i^j} + \frac{\tilde{v}^j + \tilde{v}_0^j}{K-1} \sum_{i=1}^{n} \frac{\alpha_i^j}{v_i^j}.$$

Therefore, if (59) holds, we have $\boldsymbol{\alpha} \in \mathcal{A}_{GAM}$.

Next, we show that if $\boldsymbol{\alpha} \in \mathcal{A}_{GAM}$, then (59) holds. Given $\boldsymbol{\alpha} \in \mathcal{A}_{GAM}$, define

$$w_i^j = \frac{\alpha_i^j}{p^j v_i^j} \text{ for each } i \in \mathcal{N} \text{ and } j \in \mathcal{M}, \text{ and } w_0^j = z^j = \frac{1}{\tilde{v}^j + \tilde{v}_0^j} \left(1 - \sum_{i=1}^{n} \frac{\tilde{v}_i^j \alpha_i^j}{p^j v_i^j}\right).$$

To show (27), it suffices to show the first, second and fourth inequalities hold because the rest of the constraints hold trivially.

Since $p_j \geq \sum_{i=1}^{n} \frac{\tilde{v}_i^j \alpha_i^j}{v_i^j} + \frac{\tilde{v}^j + \tilde{v}_0^j}{K-1} \sum_{i=1}^{n} \frac{\alpha_i^j}{v_i^j}$ for all j, we have

$$\sum_{i=1}^{n} w_i^j = \sum_{i=1}^{n} \frac{\alpha_i^j}{p^j v_i^j} = \frac{1}{p^j} \sum_{i=1}^{n} \frac{\alpha_i^j}{v_i^j} \leq \frac{K-1}{\tilde{v}^j + \tilde{v}_0^j} \left(1 - \sum_{i=1}^{n} \frac{\tilde{v}_i^j \alpha_i^j}{p^j v_i^j}\right) = (K-1) z^j = K z^j - w_0^j \text{ for all } j.$$

Hence, the second inequality of (59) holds. Since $p_j \geq \sum_{i'=1}^{n} \frac{\tilde{v}_{i'}^j \alpha_{i'}^j}{v_{i'}^j} + \frac{(\tilde{v}^j + \tilde{v}_0^j)\alpha_i^j}{v_i^j}$ for any i, j, we have

$$w_i^j = \frac{\alpha_i^j}{p^j v_i^j} \leq \frac{1}{\tilde{v}^j + \tilde{v}_0^j} \left(1 - \sum_{i=1}^{n} \frac{\tilde{v}_i^j \alpha_i^j}{p^j v_i^j}\right) = z^j \text{ for all } i \in \mathcal{N}, j \in \mathcal{M}.$$

Therefore, (59) holds. Hence, the feasible region of $\boldsymbol{\alpha}$ is characterized by (57). This completes the proof. □

G. Implementation Details of the Re-Solving Benchmarks

In this section, we provide the implementation details of the re-solving benchmarks: the FA-R and FA-I-R policies. For both policies, we preset a set of re-solving epochs $\mathcal{T} := \{t_u : u = 0, 1, 2, ..., U\}$ in which the algorithm re-solves the FA convex program with updated ad budgets and click-through requirements, where $t_0 = 1$ refers to solving $\mathcal{FA}(\gamma)$ at the beginning of the planning horizon. At each re-solving epoch t_u ($1 \leq u \leq U$), an FA-R or FA-I-R policy will re-solve the following convex program with budget and click-through requirement updates (similar to Appendix B, we define $Y_i^j(t) := \sum_{\tau=1}^{t-1} y_i^j(\tau)$ as the cumulative click-throughs until the beginning of time t):

$$\begin{aligned}
\max_{\boldsymbol{z}} \ & \mathcal{V}_{\mathsf{FA}}(\boldsymbol{z}|u) := \sum_{i\in\bar{\mathcal{N}}, j\in\mathcal{M}, S\in\mathfrak{S}^j} r_i^j p^j \phi_i^j(S) z^j(S) + \lambda F(\boldsymbol{\zeta}) \\
\text{s.t.} \ & \sum_{j\in\mathcal{M}, S\in\mathfrak{S}^j} b_i p^j \phi_i^j(S) z^j(S) \leq \frac{B_i - b_i \sum_{j\in\mathcal{L}_i} Y_i^j(t_u)}{T - t_u + 1} \text{ for all } i \in \mathcal{N} \\
& \sum_{j\in\mathcal{C}, S\in\mathfrak{S}^j} p^j \phi_i^j(S) z^j(S) \geq \frac{\eta_i^{\mathcal{C}} - \sum_{j\in\mathcal{C}} Y_i^j(t_u)}{T - t_u + 1} \text{ for all } i \in \bar{\mathcal{N}} \text{ and } \mathcal{C} \in \mathfrak{K}_i \\
& \sum_{S\in\mathfrak{S}^j} z^j(S) \leq 1 \text{ for all } j \in \mathcal{M} \\
& z^j(S) \geq 0 \text{ for all } j \in \mathcal{M}, S \in \mathfrak{S}^j \\
& \boldsymbol{\zeta} \in \mathbb{R}^{(n+1)m}, \text{ with } \zeta_i^j = \frac{t_u - 1}{T} \cdot Y_i^j(t_u) + \frac{T - t_u + 1}{T} \cdot \sum_{i\in S\in\mathfrak{S}^j} p^j z^j(S) \phi_i^j(S).
\end{aligned} \tag{64}$$

It is clear from the formulation that (64) is reduced to $\mathcal{FA}(\gamma)$ with $u = 0$. We denote the solution to (64) in re-solving epoch u as $\boldsymbol{z}^*(u)$. Therefore, the FA-R and FA-I-R policies re-solve (64) at the re-solving epochs $\mathcal{T}$, and follows the stationary history-independent policy $\tilde{\pi}_{\mathsf{FA}}(\boldsymbol{z}^*(u))$ from time t_u to time $t_{u+1} - 1$ for $u = 1, 2, ..., U$, where we adopt the convention $t_{U+1} = T + 1$.

The only difference between the FA-R and FA-I-R policies is the pattern of the re-solving epochs $\mathcal{T}$. More specifically, for the FA-R policy (see, also, Jasin and Kumar 2012), $\mathcal{T}$ is evenly spread across all time, i.e., $t_u = \lceil \frac{Tu}{U} \rceil$, where $\lceil \cdot \rceil$ refers to the ceiling function. For the FA-I-R policy (see, also, Bumpensanti and Wang 2020), $\mathcal{T}$ is sparser at the beginning of the planning horizon and denser at the end. Specifically, following Bumpensanti and Wang (2020), we set the re-solving epoch $t_u = \lceil T - T^{(5/6)^u} \rceil$ for all $u \in \{0, 1, 2, ..., U\}$, where $U = \lceil \frac{\log(\log(T))}{\log(6/5)} \rceil$. Finally, we remark that, in our numerical experiments, we set $T = 1{,}000$ and, therefore, $U = 7$ for the FA-I-R policy. To make the benchmark policies comparable, we set $U = 7$ for the FA-R policy as well.

H. Mean-Reverting Behavior of the DWO Policy

To highlight the mean-reverting property of our proposed DWO algorithm, we examine the intertemporal correlation between of the click-through $y_i^j(t)$ of ad i by type-j customers in period t and the *per-period debt*, defined as $\Delta_i^j(t) := d_i^j(t)/t$ where $d_i^j(t)$ is the debt of ad i for customer segment j at the beginning of time t (as defined in Algorithm 1). Recall that debt measures the gap between the click-through target set by the algorithm and the realized click-throughs. Therefore, if the correlation between $y_i^j(t)$ and $\Delta_i^j(t)$ is larger, it implies that the algorithm "pays back" the "debt" faster and, therefore, the mean-reversion of the click-through process is stronger.

For each of the 4 algorithms studied in our numerical experiments, we regress the click-through on the per-period debt using the following model specification with 30 million randomly drawn samples for each policy studied:

$$y_i^j(t) = a_0 + a_1 \Delta_i^j(t) + \epsilon$$

The regression results are reported in Table 3. If we instead regress the click-through on the total debt $d_i^j(t)$, the results will be similar because the per-period debt is a constant multiplication of the total debt.

Policy	Coefficient	Estimation	Standard Error	t-statistics	p-value
FA	a_0	0.0032852	1.0448e-05	314.44	0
	a_1	0.022207	0.0020471	10.848	2.04e-27
FA-R	a_0	0.0033432	1.0539e-05	317.23	0
	a_1	0.089427	0.0021384	41.819	0
FA-I-R	a_0	0.0032908	1.0456e-05	314.72	0
	a_1	0.0096747	0.0020433	4.7348	2.1925e-06
DWO	a_0	0.0034613	1.0635e-05	325.45	0
	a_1	0.45608	0.0026211	174	0

Table 3 The Regression Results of the Intertemporal Correlations Between Click-Throughs and Per-Period Debts

Table 3 shows that our DWO algorithm clearly drives the mean-reverting pattern for the click-through process, captured by the fact that the estimate $\hat{a}_1 = 0.456$ is positive, large and statistically significant. This is expected given that the DWO policy gives a higher weight for the ad/customer pair with a larger debt at each time t. An important observation from our regression results is that, the estimated coefficient $\hat{a}_1$ of our DWO algorithm (0.456) is about one order of magnitude larger than that of other benchmarks except for FA-R. Such an observation delivers an intriguing insight that our debt-based algorithm drives the click-through process toward its mean (i.e., the target set by the first-stage optimization) and, as a consequence, result in a more stable budget depletion process for the ads. Finally, we remark that, because of the budget constraints of the ads, the FA-based benchmarks also exhibit certain mean-reverting property weaker than our DWO algorithm.

I. Comparison With the Inventory-Balancing Policies

In this section, we compare our DWO algorithm against another family of benchmarks called the inventory balancing (IB) policies. Specifically, Golrezaei et al. (2014) propose two IB algorithms which implement real-time personalized offer-set optimization with an exponential penalty function (the EIB policy) and a linear penalty function (the LIB policy), respectively, to reweight the value of each ad. Upon the arrival of customer t, the IB policies solve a single-period offer-set optimization problem with a discounted value $r_i\Phi(B_i(t-1)/B_i)$, where $\Phi(\cdot)$ is an increasing discount function and $B_i(t-1)$ is the budget of ad i at the end of time $t-1$. The discount function is $\Phi(x) = (e/(e-1)) \cdot (1-e^{-x})$ under the EIB policy and is $\Phi(x) = x$ under the LIB policy. It is hard, if not impossible, to incorporate the non linear fairness metric and the click-through requirements into the IB policies, so we remove these modeling features in the comparison between DWO and IB policies. We consider the same numerical setup as Section 6 with the click-through requirements and the fairness metric removed.

CP	LF	EIB	LIB	DWO
0.1	1.5	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.25	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.00	99.44% (0.18%)	99.44% (0.25%)	99.20% (0.48%)
	0.75	91.79% (0.21%)	90.66% (0.28%)	99.00% (0.63%)
	0.5	89.19% (0.31%)	86.72% (0.32%)	98.96% (0.49%)
1	1.5	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.25	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.00	99.43% (0.31%)	99.45% (0.24%)	99.40% (0.39%)
	0.75	91.75% (0.17%)	90.60% (0.24%)	98.81% (0.63%)
	0.5	89.23% (0.29%)	86.86% (0.24%)	98.67% (0.64%)
10	1.5	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.25	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.00	99.54% (0.21%)	99.52% (0.27%)	99.37% (0.39%)
	0.75	91.64% (0.17%)	90.53% (0.25%)	99.19% (0.41%)
	0.5	89.20% (0.23%)	86.76% (0.24%)	98.75% (0.51%)
100	1.5	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.25	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.00	99.59% (0.25%)	99.56% (0.20%)	99.35% (0.40%)
	0.75	91.68% (0.19%)	90.53% (0.21%)	99.29% (0.29%)
	0.5	89.20% (0.18%)	86.71% (0.23%)	98.93% (0.62%)

Table 4 Numerical Results (Standard Error Relative to the Theoretical Upper Bound in Parentheses)

We report the results on the comparison between our DWO policy and the EIB and LIB algorithms in Table 4, with the ratio between the standard error of the total advertising revenue for each policy examined to the theoretical upper-bound of advertising revenue included in the parenthesis. The most important takeaway from our experiments is that the DWO policy outperforms the EIB and LIB algorithms when LF is low, especially when $LF < 1$. In this case, the budget constraints are not binding, so discounting the ad value when the budget is low is not helpful. On the other hand, when the loading factor is high, the budgets are more likely to be exhausted, so the discount functions of the EIB and LIB algorithms can help smoothly allocate the ad budgets, thus giving rise to higher efficiency performance than the DWO algorithm. To conclude this section, we remark that, because of the difficulty to incorporate the nonlinear fairness metric into the IB policies, this family of algorithms are not amenable to address the algorithmic fairness issue, which can be well handled by our DWO policy.

J. Concentration Parameter

In our numerical experiments (Section 6), we vary CP to change the uniformness of proportions $p^1, ..., p^m$ of m customer types, which are generated by a m-dimension Dirichlet distribution. The m-dimension Dirichlet distribution has m concentration parameters $\beta_1, ..., \beta_m$. In our experiments, we set $CP := \beta_0 = \beta_1 = \beta_2 = ... = \beta_m$. Note that, for all j, $\mathbb{E}[p^j] = \frac{\beta_j}{\sum_{k=1}^m \beta_k} = \frac{1}{m}$ and $Var(p^j) = \frac{m-1}{m^2(m\beta_0+1)}$, which is decreasing in β_0. For $j \neq k$, the covariance between p^j and p^k is $-\frac{1}{m^2(m\beta_0+1)}$, which is increasing in β_0. Hence, if β_0 is larger, the sampled customer type distribution will be close to the uniform distribution on $\{1, 2, ..., m\}$. In contrast, if $CP = \beta_0$ is small, the customer type distribution is more likely to be concentrated on a subset of $\{1, 2, ..., m\}$. In other words, the higher the CP, the more uniform the generated distribution of customer types.